

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2005, 16:45:24 ; Search time 172 Seconds
(without alignments)
53.590 Million cell updates/sec

Title: US-10-031-289-1331
Perfect score: 88
Sequence: 1 PTOKAEINQSKLEEQ 18

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 259284

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Listing first 100 summaries

Database : Uniprot 03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	59.1	59	2	003122 streptococc
2	44	50.0	98	2	082XR2
3	43	48.9	75	2	054581 streptococc
4	42	47.7	91	2	0708P5 bacterioph
5	41	46.6	71	2	065MG3 bacterioph
6	41	46.6	76	2	064342 bacterioph
7	41	46.6	92	2	06XN41 rhodococcus
8	40	45.5	47	2	07VFX3 helicobacte
9	40	45.5	76	2	06UAV4 bacterioph
10	40	45.5	83	2	09XV13 vibrio chol
11	39	44.3	71	2	076B50 homo sapien
12	39	44.3	92	2	065IU0 bacillus li
13	39	44.3	92	2	054838 streptococc
14	39	44.3	92	2	063TN8 bacterioph
15	38.5	43.8	83	2	074MX1 nanosarchaeu
16	38	43.2	60	2	065HE9 bacterioph
17	38	43.2	61	2	08R855 thermoaer
18	38	43.2	65	2	07VC24 prochloroc
19	38	43.2	84	2	0832A2 enterococcu
20	38	43.2	88	2	061VZ4 photobacter
21	38	43.2	95	1	08TC MOUSE
22	38	43.2	95	2	06TMJ9 mus musculi
23	37	42.0	53	2	07VVC2 trypanosoma
24	37	42.0	67	2	097N79 streptococc
25	37	42.0	67	2	08CY60 streptococc
26	37	42.0	70	2	046572 equus cabal
27	37	42.0	75	2	054665 streptococc
28	37	42.0	79	2	054642 streptococc
29	37	42.0	79	2	054649 streptococc
30	37	42.0	79	2	054660 streptococc
31	37	42.0	80	2	090QH4 human immun

32	37	42.0	81	2	098PH3 mycoplasma
33	37	42.0	83	2	054643 streptococc
34	37	42.0	83	2	087TN3 vibrio para
35	37	42.0	88	2	054588 streptococc
36	37	42.0	92	2	054644 streptococc
37	37	42.0	92	2	09AQ58 bacterioph
38	37	42.0	94	2	08PT46 methanobarc
39	37	42.0	94	2	054540 streptococc
40	37	42.0	96	2	054583 streptococc
41	37	42.0	96	2	091443 streptococc
42	37	42.0	96	2	08DXH1 xenopus lae
43	37	42.0	96	2	098SV3 mycoplasma
44	36.5	41.5	97	2	06MR78 plasmodium
45	36.5	41.5	97	2	0813G7 myxococcus
46	36	40.9	52	2	08KRC2 homo sapien
47	36	40.9	54	2	08DXH1 streptococc
48	36	40.9	67	1	RL29_BACHD
49	36	40.9	68	2	093T87 enterococcu
50	36	40.9	71	2	079DF7 caenorhabdi
51	36	40.9	72	2	018168 bartonella
52	36	40.9	73	2	06G207 bacterioph
53	36	40.9	74	2	09XJ55 streptococc
54	36	40.9	74	2	003117 streptococc
55	36	40.9	75	2	054553 streptococc
56	36	40.9	82	2	0832C3 enterococcu
57	36	40.9	91	2	08KPO2 synechococc
58	36	40.9	92	2	0616B3 escherichia
59	36	40.9	93	2	091TV9 caenorhabdi
60	36	40.9	95	2	08R893 thermoaer
61	36	40.9	95	2	08K2Y1 mus musculi
62	36	40.9	96	2	08NKV7 acidianus a
63	36	40.9	96	2	032415 pseudomonas
64	36	40.9	96	2	07P280 fusbacteri
65	36	40.9	96	2	0886K7 pseudomonas
66	36	40.9	97	2	09P0G1 homo sapien
67	36	40.9	98	2	09ZH16 desulfotoma
68	35	39.8	27	2	016162 homo sapien
69	35	39.8	36	2	009095 avena sativ
70	35	39.8	48	2	08E517 streptococc
71	35	39.8	53	2	08BD10 thermoaer
72	35	39.8	61	2	09ANN0 bradyrhizob
73	35	39.8	63	2	081F83 trypanosoma
74	35	39.8	72	2	065J36 bacillus li
75	35	39.8	75	1	EX75_THEMA
76	35	39.8	76	2	072B12 desulfovibr
77	35	39.8	76	2	08DXK9 streptococc
78	35	39.8	78	2	09KSN3 vibrio chol
79	35	39.8	79	2	054604 streptococc
80	35	39.8	80	2	084933 streptococc
81	35	39.8	82	2	080Y75 passion fru
82	35	39.8	83	2	09S1K7 streptococc
83	35	39.8	83	2	07M0H2 vibrio vuln
84	35	39.8	83	2	08DDC7 vibrio vuln
85	35	39.8	86	2	081P74 bacillus an
86	35	39.8	86	2	06HHG8 bacillus th
87	35	39.8	86	2	09DYE0 human herpe
88	35	39.8	86	2	09WT47 human herpe
89	35	39.8	87	2	09CW88 mus musculi
90	35	39.8	91	2	065K08 bacillus li
91	35	39.8	94	2	062W55 bacillus li
92	35	39.8	96	2	028610 archaeoglob
93	35	39.8	96	2	021177 caenorhabdi
94	35	39.8	96	2	09MAR2 staphylococ
95	35	39.8	96	2	09AE02 pseudomonas
96	35	39.8	96	2	09AQ02 pseudomonas
97	34	38.6	34	2	07R6J6 giardia lam
98	34	38.6	38	2	06MMJ4 bdellovibri
99	34	38.6	38	2	091VC8 mus musculi
100	34	38.6	40	2	08NWX0 staphylococ

ALIGNMENTS

RESULT 1

003122 PRELIMINARY; PRT; 59 AA.
AC 003122;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE M-like protein (Fragment).
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93084815; PubMed=1339461;
RA Reif W.A., Martin D.R., Striprakash K.S.;
RT "Identification of sequence types among the M-nontypeable group A
RT streptococci [see comments]."
RL J. Clin. Microbiol. 30:3190-3194(1992).
DR EMBL, L05024; AAA21790.1; -.
FT NON_TER 1
FT TER 59
SQ SEQUENCE 59 AA; 6996 MW; FA7A45ADAJA26857 CRC64;

Query Match 59.1%; Score 52; DB 2; Length 59;
Best Local Similarity 64.7%; Pred. No. 3.2;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PTKAAELNOKSKELBOQ 17
|:|||||:|
Db 7 PVKKAELYDKIKLELE 23

RESULT 2

082XR2 PRELIMINARY; PRT; 98 AA.
AC 082XR2;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE - Hypothetical protein.
GN OrderedLocusNames=NE0192;
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19718 / IFO 14298;
RX MEDLINE=22586410; PubMed=12700255;
DOI=10.1128/JB.185.9.2759-2773.2003;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.L.,
RA Hauer L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA Arciero D.M., Holmes N.G., Whitaker M.M., Arp D.J.;
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
RT obligate chemolithoautotroph Nitrosomonas europaea."
RL J. Bacteriol. 185:2759-2773(2003).
DR EMBL, BX321856; CAD84103.1; -.
DR InterPro; IPR007475; DUF526.
DR Pfam; PF04380; DUF526; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 98 AA; 11073 MW; BCC9650CA16C63ED CRC64;

Query Match 50.0%; Score 44; DB 2; Length 98;
Best Local Similarity 60.0%; Pred. No. 81;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 4 KAAELNOKSKELBOQ 18
|:|:|:|:|:|:|
Db 62 KLAELBEKVRKLEBOQ 76

RESULT 3

054581 PRELIMINARY; PRT; 75 AA.
AC 054581;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Emml protein (Fragment).
GN Name=emml;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M type 65;
RX MEDLINE=95188537; PubMed=7891551;
RA Whatmore A.M., Kapur V., Sullivan D.J., Musser J.M., Kehoe M.A.;
RT "Non-congruent relationships between variation in emm gene sequences
RT and the population genetic structure of group A streptococci."
RL Mol. Microbiol. 14:619-631(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=M type 65;
RA Whatmore A.M.;
RT "Sequence Analysis of the Emm-like Gene Family of Streptococcus
RT Progenes."
RL Thesis (1993), Microbiology, University of Newcastle Upon Tyne.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=M type 65;
RA Whatmore A.M.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL, U11980; AAA95956.1; -.
DR PIR; S60849; S60849. 1
FT NON_TER 1
FT TER 75
SQ SEQUENCE 75 AA; 8311 MW; 6DFC661CAE097FE6 CRC64;

Query Match 48.9%; Score 43; DB 2; Length 75;
Best Local Similarity 47.1%; Pred. No. 86;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TOKAAELNOKSKELBOQ 18
|:|:|:|:|:|:|
Db 56 TEKNEBLDKKXKGLDSQ 72

RESULT 4

0708P5 PRELIMINARY; PRT; 91 AA.
AC 0708P5;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Bacteriophage EJ-1.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=12402;
RN [1]
RP SEQUENCE FROM N.A.
RC PubMed=15110522; DOI=10.1016/j.virol.2004.01.029;
RX Romero P., Lopez R., Garcia B.;
RT "Genomic organization and molecular analysis of the inducible prophage
RT EJ-1, a mosaic myovirus from an atypical pneumococcus."
RL Virology 322:239-252(2004).
RN [2]
RP SEQUENCE FROM N.A.
RA Diaz E., Lopez R., Garcia J.L.;
RT "EJ-1, a temperate bacteriophage of Streptococcus pneumoniae with a
RT Myoviridae morphotype."
RL J. Bacteriol. 174:5516-5525(1992).
DR EMBL, AJ609634; CAE82110.1; -.

KM Hypothetical protein.
SQ SEQUENCE 91 AA; 10594 MW; 1501CB9209FE3BC0 CRC64;

Query Match 47.7%; Score 42; DB 2; Length 91;
Best Local Similarity 53.3%; Pred. No. 1.5e+02;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 OKAAELNOKSKELEQ 17
:|||||:|:|:|:
DB 7 EKAGEIRQIKLEK 21

RESULT 5

065MG3 PRELIMINARY; PRT; 71 AA.

AC 065MG3;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein yfkK.
GN Name=yfkK; ORFNames=BL03111, BL100816;
OS Bacillus licheniformis DSM 13.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 13;
RX PubMed=15383718;
RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baumeier S., Henne A., Liesegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
RT Organism with Great Industrial Potential.";
RL J. Mol. Microbiol. Biotechnol. 7:204-211 (2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14580;
RX
RA Rey M.W., Ranaiva P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang M., de Leon A.L., Xiang H., Gucsi V., Clausen I.G., Olsen P.B.,
RA Rasmussen M.D., Andersen J.T., Jorgensen P.L., Larsen T.S.,
RA Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D.,
RA Betka R.M.;
RT "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species.";
RL Genome Biol. 5:R77-R77(2004).
DR EMBL; AB017333; AAU39751.1; -;
DR EMBL; CP000002; AAU22403.1; -;
KW Hypothetical protein.
SQ SEQUENCE 71 AA; 8188 MW; 24A55A92709E20C3 CRC64;

Query Match 46.6%; Score 41; DB 2; Length 71;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 OKAAELNOKSKELEQ 18
:|||||:|:|:|:
DB 4 OKTAEINMIEISOK 19

RESULT 6

064342 PRELIMINARY; PRT; 76 AA.

AC 064342;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Gp1.
GN Name=Gene 31;
OS Bacteriophage N15.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=40631;
RN [1]

RP SEQUENCE FROM N.A.
RA Hendrix R.W., Ravin V.K., Caejens S.R., Ford M.E., Ravin N.V.,
RA Smirnov I.K.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF064539; AACL9073.1; -;
DR PIR; T13117; T13117.
SQ SEQUENCE 76 AA; 8373 MW; 32C0831B90C4B1A6 CRC64;

Query Match 46.6%; Score 41; DB 2; Length 76;
Best Local Similarity 56.2%; Pred. No. 1.7e+02;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 OKAAELNOKSKELEQ 18
:|||||:|:|:|:
DB 43 OKAAELIQVARELTQE 58

RESULT 7

06XN41 PRELIMINARY; PRT; 92 AA.

AC 06XN41;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Putative transposase.
GN ORFNames=BD2.105;
OS Rhodococcus erythropolis.
OC Plasmid pBD2.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=1833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BD2;
RX MEDLINE=22803265; PubMed=12923100;
RA DOI=10.1128/JB.185.17.5269-5274.2003;
RA Stecker C., Johann A., Herzberg C., Averhoff B., Gottschalk G.;
RT "Complete nucleotide sequence and genetic organization of the 210-
RT kilobase linear plasmid of Rhodococcus erythropolis BD2.";
RL J. Bacteriol. 185:5269-5274(2003).
DR EMBL; AY223810; AAP73990.1; -;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004803; F:transposase activity; IEA.
DR GO; GO:0006313; P:DNA transposition; IEA.
KW Plasmid.
SQ SEQUENCE 92 AA; 10732 MW; 3466368BAE7E816 CRC64;

Query Match 46.6%; Score 41; DB 2; Length 92;
Best Local Similarity 41.2%; Pred. No. 2.1e+02;
Matches 7; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 2 TOKAAELNOKSKELEQ 18
:|||||:|:|:|:
DB 50 SEKNAEASKRTKELEKE 66

RESULT 8

Q7VEY3 PRELIMINARY; PRT; 47 AA.

AC Q7VEY3;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=HH1542;
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 51449 / 381;
RX MEDLINE=22709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;

```

RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Briand P.,
RA Bell M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Froesch M., Fox J.G.,
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906 (2003).
KM EMBL: AB017148; AAF78139.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 47 AA; 5646 MW; 08156C320356548E CRC64;

Query Match 45.5%; Score 40; DB 2; Length 47;
Best Local Similarity 41.2%; Pred. No. 1.5e+02;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 2 TOKAELNOKSKELEEQ 18
Db 23 SERKAELERKIREESQ 39

RESULT 9
Q6UAV4 PRELIMINARY; PRT; 76 AA.
AC O6UAV4;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
GN ORFNames=28;
OS Bacteriophage phiKO2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
CX NCBI_TaxID=255431;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14996813; DOI=10.1128/JB.186.6.1818-1832.2004;
RA Casjens S.R., Gilcrease E.B., Huang W.M., Buny K.L., Pedulla M.L.,
RA Ford M.B., Houtz J.M., Hatfull G.F., Hendrix R.W.;
RT "The pKO2 linear plasmid prophage of Klebsiella oxytoca."
RL J. Bacteriol. 186:1818-1832 (2004).
RN [2]
RP SEQUENCE FROM N.A.
RA Casjens S., Pedulla M.L., Ford M.B., Houtz J.M., Gilcrease E.B.,
RA Huang W.M., Buny K.L., Hatfull G.F., Hendrix R.W.;
RL Submitted (AUG-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL: AY344448; AAR83044.1; -.
SQ SEQUENCE 76 AA; 8381 MW; 47435C722518F0E2 CRC64;

Query Match 45.5%; Score 40; DB 2; Length 76;
Best Local Similarity 56.2%; Pred. No. 2.4e+02;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 OKAELNOKSKELEEQ 18
Db 43 QEAELLOVARELVOE 58

RESULT 10
Q9KV13 PRELIMINARY; PRT; 83 AA.
AC Q9KV13;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein VC0163;
GN OrderedLocustNames=VC0163;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
CX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406633; PubMed=10952301; DOI=10.1038/35020000;

```

```

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwim M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Esmolaeva M.D., Vamathevan U.J., Bass S., Qin H., Dreogl I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483 (2000).
DR EMBL: AB004106; AAF93339.1; -.
DR PIR: H82356; H82356.
DR TIGR: VC0163; -.
DR InterPro: IPR007475; DUF526.
DR Pfam: PF04380; DUF526; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 83 AA; 9600 MW; 241BA0D6B7AED6E CRC64;

Query Match 45.5%; Score 40; DB 2; Length 83;
Best Local Similarity 64.3%; Pred. No. 2.7e+02;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 OKAELNOKSKELE 16
Db 64 OKLELOKMAELE 77

RESULT 11
Q76B50 PRELIMINARY; PRT; 71 AA.
AC Q76B50;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE T1560.
GN Name=T1560;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hishinuma A., Ohmika N., Namatame T., Ielri T.;
RT "TTF-2 stimulates expression of 17 genes, including one novel thyroid-
RT specific gene which might be involved in thyroid development."
RL Mol. Cell. Endocrinol. 221:33-46 (2004).
DR EMBL: AB111913; BAD04069.1; -.
SQ SEQUENCE 71 AA; 8156 MW; BC81CB37EB378292 CRC64;

Query Match 44.3%; Score 39; DB 2; Length 71;
Best Local Similarity 46.7%; Pred. No. 3.2e+02;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TOKAELNOKSKELE 16
Db 11 TRKEIQVNOKEOKLE 25

RESULT 12
Q6S1U0 PRELIMINARY; PRT; 74 AA.
AC Q6S1U0;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Yxh (Hypothetical protein).
GN Name=Yxh; ORFNames=BL05199; BL102137;
OS Bacillus licheniformis DSM 13.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_TaxID=279010;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 13;

```

RX PubMed=15383718;
 RA Velth B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
 RA Ehrenreich P., Baumeier S., Henne A., Liesegang H., Mewli R.,
 RA Ehrenreich A., Gottschalk G.;
 RT "The Complete Genome Sequence of *Bacillus licheniformis* DSM13, an
 RT Organism with Great Industrial Potential.";
 RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 14580;
 RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
 RA Tang M., de Leon A.L., Xiang V., Gusti V., Clausen I.G., Olsen P.B.,
 RA Rasmussen M.D., Andersen J.T., Jorgensen P.L., Larsen T.S.,
 RA Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D.,
 RA Berka R.M.;
 RT "Complete genome sequence of the industrial bacterium *Bacillus*
 RT *licheniformis* and comparisons with closely related *Bacillus* species.";
 RL Genome Biol. 5:R77-R77(2004).
 DR EMBL; AE017333; AAU41024.1; -;
 DR EMBL; CP000002; AAU23662.1; -;
 KM Hypothetical protein.
 SQ SEQUENCE 74 AA; 8473 MW; 64A61F6A98030071 CRC64;
 Query Match 44.3%; Score 39; DB 2; Length 74;
 Best Local Similarity 43.8%; Pred. No. 3.3e+02;
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 Qy 3 QKAAELNOKSKLEEQ 18
 Db 22 QSSQOMNORLKNLEQ 37
 RESULT 13
 Q54838 PRELIMINARY; PRT; 92 AA.
 ID Q54838;
 AC Q54838;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE M42 protein (Fragment).
 DE Name=emm42;
 GN Streptococcus pyogenes.
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 CX NCBI_TaxID=1314;
 CX [1]
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=M42;
 RA Beall B.W.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U46799; AAA80970.1; -;
 FT NON TER 1 1
 FT NON TER 92 92
 SQ SEQUENCE 92 AA; 10252 MW; 43A251B474BAE96 CRC64;
 Query Match 44.3%; Score 39; DB 2; Length 92;
 Best Local Similarity 43.8%; Pred. No. 4.1e+02;
 Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
 Qy 3 QKAAELNOKSKLEEQ 18
 Db 70 EKNEBLDKKKNLEDK 85
 RESULT 14
 Q63TN8 PRELIMINARY; PRT; 92 AA.
 ID Q63TN8;
 AC Q63TN8;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Putative exported protein.
 GN ORFNames=BPSL1929;

OS Burkholderia pseudomallei K96243.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia.
 CX NCBI_TaxID=272560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K96243;
 RX PubMed=15377794;
 RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.M.,
 RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
 RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
 RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
 RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshazer D.,
 RA Feltwell T., Fraser A., Hance Z., Hauser H., Holtroyd S., Jagsels K.,
 RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
 RA Rabinowitsch E., Rutherford K., Sanders M., Simmonds M.,
 RA Songsvilal S., Stevens K., Tumapa S., Vesaratchavee M.,
 RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
 RT "Genomic plasticity of the causative agent of melioidosis,
 RT *Burkholderia pseudomallei*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
 DR EMBL; BX571965; CAH35928.1; -;
 SQ SEQUENCE 92 AA; 9649 MW; 419F2FCF3654CFED CRC64;
 Query Match 44.3%; Score 39; DB 2; Length 92;
 Best Local Similarity 46.7%; Pred. No. 4.1e+02;
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 Qy 4 KAAELNOKSKLEEQ 18
 Db 35 KIAIEQRAQALQOQ 49
 RESULT 15
 Q74MX1 PRELIMINARY; PRT; 83 AA.
 ID Q74MX1;
 AC Q74MX1;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE NEQ516.
 GN OrderedLocustNames=NEQ516;
 OS Nanoarchaeum equitans.
 OC Archaea; Nanoarchaeota; Nanoarchaeum.
 CX NCBI_TaxID=160232;
 CX [1]
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=K14-M;
 RX MEDLINE=22946215; PubMed=14566062; DOI=10.1073/pnas.1735403100;
 RA Waters E., Hohn M.J., Abel I., Graham D.E., Adams M.D., Barnstead M.,
 RA Beeson K.Y., Bibbs U., Bolanos R., Keller M., Kretz K., Lin X.,
 RA Mathur E., Ni J., Podar M., Richardson T., Sutton G.G., Simon M.,
 RA Soell D., Stetter K.O., Short J.M., Noordermeer M.;
 RT "The genome of *Nanoarchaeum equitans*: insights into early archaeal
 RT evolution and derived parasitism.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:12984-12988(2003).
 CC -1- Similarity: Belongs to the prefolidin beta subunit family.
 DR EMBL; AE017199; AAR39357.1; -;
 DR InterPro; IPR002777; PFD_beta_1like.
 DR InterPro; IPR009053; Prefoldin.
 DR Pfam; PF01920; KE2; 1.
 KM Chapterone; Complete proteome.
 SQ SEQUENCE 83 AA; 9725 MW; 871BC6E874661BFB CRC64;
 Query Match 43.8%; Score 38.5; DB 2; Length 83;
 Best Local Similarity 58.8%; Pred. No. 4.4e+02;
 Matches 10; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
 Qy 1 PTKA-AELNOKSKLE 16
 Db 51 PKKALEELNKKKLE 67

```

RESULT 16
Q65HE9 PRELIMINARY; PRT; 60 AA.
AC Q65HE9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein ygzE
GN Name=ygzE; ORFNames=BL05265, BL102641;
OS Bacillus licheniformis DSM 13.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 13;
RX PubMed=15383718;
RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baumeier S., Henne A., Liesegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
RT Organism with Great Industrial Potential.";
RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14580;
RA Ray M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang M., de Leon A.L., Xiang H., Gusti V., Clausen I.G., Olsen P.B.,
RA Rasmussen M.D., Andersen J.T., Jorgensen P.L., Larsen T.S.,
RA Sorokin A., Bolotin A., Lapidus A., Gallerton N., Ehrlich S.D.,
RA Berka R.M.;
RT "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species.";
RL Genome Biol. 5:R77-R77(2004).
DR EMBL; AE017333; AAU4151.1; -.
DR EMBL; CP000002; AAU24155.1; -.
KM Hypothetical protein.
SQ SEQUENCE 60 AA; 7459 MW; A468ED129E4E54D CRC64;

Query Match 43.2%; Score 38; DB 2; Length 60;
Best Local Similarity 43.8%; Pred. No. 3.7e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 PTOKAEINQSKKELEQ 16
Db 21 PKDERKMKERKRELK 36

RESULT 17
Q8R855 PRELIMINARY; PRT; 61 AA.
AC Q8R855;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=TT2173;
OS Thermanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermanaerobacteriales;
OC Thermanaerobacteriaceae; Thermanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4;
RX MEDLINE=21992816; PubMed=11997336; DOI=10.1101/gr.219302;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of the T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; AE013164; AAM25332.1; -.
DR InterPro; IPR010982; Lambda_1like_DNA.
KM Complete proteome.
SQ SEQUENCE 61 AA; 7096 MW; 1F7B325AB545F606 CRC64;

```

```

Query Match 43.2%; Score 38; DB 2; Length 61;
Best Local Similarity 57.1%; Pred. No. 3.8e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 4 KAEINQSKKELEQ 17
Db 16 RLAEINMQKELSQ 29

RESULT 18
Q7VCZ4 PRELIMINARY; PRT; 65 AA.
AC Q7VCZ4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Predicted protein.
GN OrderedLocusNames=Pro0595;
OS Prochlorococcus marinus.
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1219;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SRG / CMP 1375 / SS120;
RX MEDLINE=22810154; PubMed=12917486; DOI=10.1073/pnas.173211100;
RA Dufresne A., Sallanoubat M., Patrensky F., Artiguenave F., Axmann I.M.,
RA Barde V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,
RA Makarova K.S., Ostrowski M., Ozias S., Robert C., Rogozin I.B.,
RA Scanlan D.J., Tandeau de Marsac N., Weissenbach J., Wincker P.,
RA Wolf Y.I., Hess W.R.;
RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
RT a nearly minimal ocyphotrophic genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).
DR EMBL; AE017162; AAP99640.1; -.
KM Complete proteome.
SQ SEQUENCE 65 AA; 7611 MW; 0FE4FC7C8DE80059 CRC64;

Query Match 43.2%; Score 38; DB 2; Length 65;
Best Local Similarity 66.7%; Pred. No. 4.1e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TOKAEINQSKK 13
Db 13 SEKAEINQSKK 24

RESULT 19
Q832A2 PRELIMINARY; PRT; 84 AA.
AC Q832A2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=EF2339;
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=2250857; PubMed=12663927; DOI=10.1126/science.1080613;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Tettelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,
RA Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
RA Khouri H.M., Uettermack T.R., Radune D., Ketchum K.A., Dougherty B.A.,
RA Frazer C.M.;
RT "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis.";

```

```

RL Science 299:2071-2074(2003) .
DR EMBL; AE016954; AA082064.1; -.
DR TIGR; EF2339; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 84 AA; 9683 MW; 802F5CC98FF45232 CRC64;

Query Match
Best Local Similarity 43.2%; Score 38; DB 2; Length 84;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 TQKAEINQSKELQ 17
Db 40 TDTREINQVKELEK 55

RESULT 20
ID Q6LVZ4 PRELIMINARY; PRT; 88 AA.
AC Q6LVZ4;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein YQIC.
GN Name=YQIC; OrderedLocustNames=BPRA0086;
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=74109;
RN [1]
RP SEQUENCE FROM N.A.
RA Verzi A., Campanaro S., D'Angelo M., Simonato F., Viculo N., Lauro F.,
RA Cattero A., Malacrida G., Simonati B., Cannata N., Bartlett D.,
RA Valle G.;
RT "Genome analysis of Photobacterium profundum reveals the complexity of
RT high pressure adaptations."
RL EMBL; CR378663; CAG18531.1; -.
DR InterPro; IPR007475; DUF526.
DR Pfam; PF04380; DUF526; 1.
KW Complete proteome.
SQ SEQUENCE 88 AA; 10352 MW; 4DD5E544CB115E6 CRC64;

Query Match
Best Local Similarity 43.2%; Score 38; DB 2; Length 88;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSELEQ 18
Db 64 QKTELEKMALEKK 79

RESULT 21
ID QSTC_MOUSE STANDARD; PRT; 95 AA.
AC P04641;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Osteocalcin precursor (Gamma-carboxyglutamic acid-containing protein)
DE (Bone Gla-protein) (BGP).
GN Name=Bglap1;
GN and
GN Name=Bglap2;
OS Mus musculus (Mouse).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87004555; PubMed=3019668;
RA Celeste A.J., Buecker J.L., Kriz R., Wang B.A., Mooney J.M.;
RT "Isolation of the human gene for bone gla protein utilizing mouse and
RT rat cDNA clones."

```

```

RL EMBO J. 5:1885-1890(1986) .
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94062692; PubMed=8243336; DOI=10.1210/en.133.6.3050;
RA Rahman S., Oberdorf A., Montecino M., Tanhauser S.M., Han J.B.,
RA Stein G.S., Lapidis P.J., Stein J.L.;
RT "Multiple copies of the bone-specific osteocalcin gene in mouse and
RT rat."
RL Endocrinology 133:3050-3053(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94117426; PubMed=8288580;
RA Desbois C., Hoque D.A., Karsenty G.;
RT "The mouse osteocalcin gene cluster contains three genes with two
RT separate spatial and temporal patterns of expression."
RL J. Biol. Chem. 269:1183-1190(1994).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Calvaria;
RA Yotov W.V., St Arnaud R.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Constitutes 1-2% of the total bone protein. It binds
CC strongly to apatite and calcium.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Bone.
CC -1- PTM: Gamma-carboxyglutamate residues are formed by vitamin K
CC dependent carboxylation. These residues are essential for the
CC binding of calcium.
CC -1- MISCELLANEOUS: There are two genes coding for osteocalcin, their
CC coding sequence only differs in the signal peptide region.
CC -1- SIMILARITY: Belongs to the osteocalcin / matrix Gla-protein
CC family.
CC -1- SIMILARITY: Contains 1 gamma-carboxy-glutamate domain (Gla)
CC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04142; CA27762.1; -.
DR EMBL; S67455; AAB29145.1; -.
DR EMBL; L24429; AAA39854.1; -.
DR EMBL; L24431; AAA39856.1; -.
DR EMBL; U1542; AAB40035.1; -.
DR PIR; B25471; B25471.
DR MGD; MGI:88156; Bglap1.
DR MGD; MGI:88157; Bglap2.
DR InterPro; IPR002384; Gla bone.
DR InterPro; IPR000294; VitrK_dep_Gla.
DR Pfam; PF00594; Gla; 1.
DR PRINTS; PR00002; GLABONE.
DR SMART; SM00069; GLA; 1.
DR PROSITE; PS00011; GLA_1; 1.
DR PROSITE; PS00998; GLA_2; 1.
KW Biomimetic; Calcium-binding; Gamma-carboxyglutamic acid;
KW Signal; Vitamin K.
FT SIGNAL 1 23 Probable.
FT PROPEP 24 49 Probable.
FT CHAIN 50 95 Osteocalcin.
FT DOMAIN 46 92 Gla.
FT MOD_RES 62 62 4-carboxyglutamate (By similarity).
FT MOD_RES 66 66 4-carboxyglutamate (By similarity).
FT MOD_RES 69 69 4-carboxyglutamate (By similarity).
FT DISULFID 68 74 By similarity.
FT VARIANT 4 5 LS -> IF (in Bglap2).
FT VARIANT 11 11 A -> T (in Bglap2).
FT CONFLICT 23 23 A -> P (in Ref. 1).
SQ SEQUENCE 95 AA; 10459 MW; D4AA61134805D9B CRC64;

```

Query Match 43.2%; Score 38; DB 1; Length 95;
 Best Local Similarity 44.4%; Pred. No. 6e+02;
 Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 PTKAEINOKSKLEEQ 18
 Db 63 PTRQEELNPACDELSDQ 80

RESULT 22

Q6TMJ9 PRELIMINARY; PRT; 95 AA.
 AC O6TMJ9
 DT 05-JUN-2004 (TREMblrel. 27, Created)
 DT 05-JUN-2004 (TREMblrel. 27, Last sequence update)
 DT 05-JUN-2004 (TREMblrel. 27, Last annotation update)
 DE Synaprobrevin.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
 OX NCBI_TaxId=44689;

RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=15075287; DOI=10.1128/EC.3.2.564-566.2004;
 RA Escalante R., Iranfar N., Sastre L., Loomis W.F.;
 RT "Identification of genes dependent on the MADS box transcription
 factor StfA in Dictyostelium discoideum development.";
 RL Eukaryotic Cell 3:564-566(2004).
 DR EMBL; AY392436; AAQ98877.1; -.
 DR DictyBase; DDB0214903; synb.
 DR InterPro; IPR001388; Synaprobrevin.
 DR PRINTS; PR00219; SYNAPROBREV.
 DR PRODOM; PD001229; Synaprobrevin; 1.
 DR PROSITE; PSS0892; V_SNAKE; 1.
 SQ SEQUENCE 95 AA; 10876 MW; 49950B35250E0B01 CRC64;

Query Match 43.2%; Score 38; DB 2; Length 95;
 Best Local Similarity 46.7%; Pred. No. 6e+02;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 OKAAELNOKSKLEEQ 17
 Db 34 OKVSELTDKSENKQ 48

RESULT 23

Q7YVC2 PRELIMINARY; PRT; 53 AA.
 AC Q7YVC2
 DT 01-OCT-2003 (TREMblrel. 25, Created)
 DT 01-OCT-2003 (TREMblrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=Tb927.2.3100;
 OS Trypanosoma brucei.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxId=5691;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22789168; PubMed=12907728; DOI=10.1093/nar/gkg673;
 RA El-Sayed N.M.A., Ghedin E., Song J., Macleod A., Brinkard F.,
 RA Lakin C., Manless D., Peterson J., Hou L., Taylor S., Tweedie A.,
 RA Bileau N., Khalak H.G., Lin X., Mason T., Hamrick L., Galer E.,
 RA Blundin G., Bartholomew D., Simpson A.J., Kaul S., Zhao H., Pal G.,
 RA Van Aken S., Uetzerback T., Haas B., Koo H.L., Umayam L., Sun B.,
 RA Gerard C., Leech V., Qi R., Zhou S., Schwartz D., Feldblum T.,
 RA Salzberg S., Tait A., Turner M.R., Ullu E., White O., Melville S.,
 RA Adams M.D., Fraser C.M., Donelson J.E.;
 RT "The sequence and analysis of Trypanosoma brucei chromosome II.";
 RL Nucleic Acids Res. 31:4856-4863(2003).
 DR EMBL; AE017168; AAQ15767.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 53 AA; 6467 MW; 7499B826C25944CC CRC64;

Query Match 42.0%; Score 37; DB 2; Length 53;
 Best Local Similarity 46.7%; Pred. No. 4.6e+02;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 OKAAELNOKSKLEEQ 17
 Db 19 KKAKKMKKOKTTEQ 33

RESULT 24

Q97N79 PRELIMINARY; PRT; 67 AA.
 AC Q97N79
 DT 01-OCT-2001 (TREMblrel. 18, Created)
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE Hypothetical protein SP2183.
 GN OrderedLocuNames=SP2183;
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxId=1313;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC BAA-334 / TIGR4;
 RX MEDLINE=11357209; PubMed=11463916; DOI=10.1126/science.1061217;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S.N., Heidelberg J.F., Deboy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn W.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtzapple E.K., Khouri H.M., Wolf A.M., Uetzerback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblum T.V., Angiuoli S.V., Dickinson T.,
 RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 pneumoniae.";
 RL Science 293:498-506(2001).
 DR EMBL; AE007506; AAK76234.1; -.
 DR PIR; A95255; A95255.
 DR PIR; H98119; H98119.
 DR TIGR; SP2183; -.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 67 AA; 7891 MW; AFA157781B17A12C CRC64;

Query Match 42.0%; Score 37; DB 2; Length 67;
 Best Local Similarity 53.8%; Pred. No. 5.9e+02;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 6 AEINOKSKLEEQ 18
 Db 38 AEIAOKREIEQR 50

RESULT 25

Q8CY60 PRELIMINARY; PRT; 67 AA.
 AC Q8CY60
 DT 01-MAR-2003 (TREMblrel. 23, Created)
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE Hypothetical protein spr1987.
 GN OrderedLocuNames=spr1987;
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxId=171101;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=1429245; PubMed=11544234;
 RX DOI=10.1128/JB.183.19.5709-5717.2001;
 RA Hoskins J., Alborn W.E., Jr., Arnold J., Blaszcak L.C., Burgett S.,
 RA DeHoff B.S., Estrem S.T., Filtz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmour R., Glaes J.S., Khoja H., Kraft A.R., Lagace R.E.,

RA Leblanc D.J., Lee L.N., Iefkowitz E.J., Lu J., Matsushima P.,
 RA McBrean S.M., McMenemy M., Mcleaster K., Mundy C.W., Nicot T.I.,
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P.,
 RA Sun P.-W., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skatrud P.L.,
 RA Glass J.I.;
 RT "Genome of the bacterium *Streptococcus pneumoniae* strain R6.";
 RL J. Bacteriol. 183:5709-5717(2001).
 DR EMBL; AE008563; AAL00789.1; -.
 DR PIR; A95255; A95255.
 DR PIR; H98119; H98119.
 KW Complete proteome.
 SQ SEQUENCE 67 AA; 7891 MW; AFA157781B17A12C CRC64;

Query Match 42.0%; Score 37; DB 2; Length 67;
 Best Local Similarity 53.8%; Pred. No. 5.9e+02;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Oy 6 AELNOKSELEQ 18
 Db 38 AELNOKSELEQ 50

RESULT 26
 ID 046572 PRELIMINARY; PRT; 70 AA.
 AC 046572;
 DT 01-JUN-1998 (TREMREL. 06, Created)
 DT 01-JUN-1998 (TREMREL. 06, Last sequence update)
 DT 01-JUN-2003 (TREMREL. 24, Last annotation update)
 DE Connexin 43 (Fragment).
 GN Name=CK43;
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Endometrial;
 RA Day W.E., Burghardt R.C., Ing N.H.;
 RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: One gap junction consists of a cluster of closely packed
 CC pairs of transmembrane channels, the connexons, through which
 CC materials of low MW diffuse from one cell to a neighboring cell
 CC (By similarity).
 CC -1- SUBUNIT: A connexon is composed of a hexamer of connexins (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the connexin family.
 DR EMBL; AF042352; AAB97428.1; -.
 DR GO; GO:0005922; C:connexon complex; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0015285; F:connexon channel activity; IEA.
 DR GO; GO:0007154; P:cell communication; IEA.
 DR InterPro: IPR000500; Connexin.
 DR Pfam; PF00029; Connexin; 1.
 DR SMART; SM00037; CNX; 1.
 DR PROSITE; PS00407; CONNEXIN_1; 1.
 KW Gap junction; Transmembrane.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 70 AA; 8144 MW; 799CA05965D2C31B CRC64;

Query Match 42.0%; Score 37; DB 2; Length 70;
 Best Local Similarity 46.7%; Pred. No. 6.2e+02;
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Oy 2 TOKAAELNOKSELE 16
 Db 47 TRKEKLNKKEELK 61

RESULT 27

OS04665
 ID 054665 PRELIMINARY; PRT; 75 AA.
 AC 054665;
 DT 01-NOV-1996 (TREMREL. 01, Created)
 DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
 DT 01-JUN-2003 (TREMREL. 24, Last annotation update)
 DE Emml protein (Fragment).
 GN Name=emml;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M type 28;
 RX MEDLINE=95349390; PubMed=7623660;
 RA Whatmore A.M., Kapur V., Musser J.M., Kehoe M.A.;
 RT "Molecular population genetic analysis of the emm subdivision of group
 RT A streptococcal emm-like genes: horizontal gene transfer and
 RT restricted variation among emm genes.";
 RL Mol. Microbiol. 15:1039-1048(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M type 44;
 RX MEDLINE=95349390; PubMed=7623660;
 RA Whatmore A.M., Kapur V., Musser J.M., Kehoe M.A.;
 RT "Molecular population genetic analysis of the emm subdivision of group
 RT A streptococcal emm-like genes: horizontal gene transfer and
 RT restricted variation among emm genes.";
 RL Mol. Microbiol. 15:1039-1048(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M type 44;
 RA Whatmore A.M.;
 RT "Sequence Analysis of the Emm-Like Gene Family of Streptococcus
 RT Pyogenes.";
 RL Thesis (1993), Microbiology, University of Newcastle Upon Tyne.
 DR EMBL; U20826; AAA87907.1; -.

SQ SEQUENCE 75 AA; 8485 MW; EDC004480628B73 CRC64;

Query Match 42.0%; Score 37; DB 2; Length 75;
 Best Local Similarity 50.0%; Pred. No. 6.6e+02;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 3 OKAAELNOKSELEQ 18
 Db 45 KLEBOKSKNLEQ 60

RESULT 28
 ID 054642 PRELIMINARY; PRT; 79 AA.
 AC 054642;
 DT 01-NOV-1996 (TREMREL. 01, Created)
 DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
 DT 01-JUN-2003 (TREMREL. 24, Last annotation update)
 DE Emml protein (Fragment).
 GN Name=emml;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M type 44;
 RX MEDLINE=95349390; PubMed=7623660;
 RA Whatmore A.M., Kapur V., Musser J.M., Kehoe M.A.;
 RT "Molecular population genetic analysis of the emm subdivision of group
 RT A streptococcal emm-like genes: horizontal gene transfer and
 RT restricted variation among emm genes.";
 RL Mol. Microbiol. 15:1039-1048(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M type 44;
 RA Whatmore A.M.;
 RT "Sequence Analysis of the Emm-Like Gene Family of Streptococcus
 RT Pyogenes.";
 RL Thesis (1993), Microbiology, University of Newcastle Upon Tyne.
 DR EMBL; U20826; AAA87907.1; -.

```

DR   PIR; S62072; S62072. 1
FT   NON_TER 1
RT   NON_TER 79
SQ   SEQUENCE 79 AA; 8687 MW; 296DCC0A12104B81 CRC64;

Query Match
Best Local Similarity 42.0%; Score 37; DB 2; Length 79;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QKAAELNQSKELEQQ 18
   :|:|:|:|:|:|
Db 55 KKLSEOEKSKNLEKQ 70

RESULT 29
Q54649 PRELIMINARY; PRT; 79 AA.
ID   Q54649;
AC   Q54649;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Emml protein (Fragment).
GN   Name=emml;
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   STRAIN=M type 15;
RC   STRAIN=M type 15;
RA   Whatmore A.M.; Kapur V.; Musser J.M.; Kehoe M.A.;
RT   "Molecular population genetic analysis of the emm subdivision of group
RT   A streptococcal emm-like genes: horizontal gene transfer and
RT   restricted variation among emm genes.";
RL   Mol. Microbiol. 15:1039-1048(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M type 15;
RA   Whatmore A.M.;
RT   "Sequence Analysis of the Emm-Like Gene Family of Streptococcus
RT   pyogenes.";
RL   Thesis (1993), Microbiology, University of Newcastle Upon Tyne.
DR   EMBL; U20833; AAA87914.1; -.
DR   PIR; S61802; S61802. 1
FT   NON_TER 1
RT   NON_TER 79
SQ   SEQUENCE 79 AA; 8829 MW; A222346A82572E4A CRC64;

Query Match
Best Local Similarity 42.0%; Score 37; DB 2; Length 79;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QKAAELNQSKELEQQ 18
   :|:|:|:|:|:|
Db 51 KKLSEOEKSKNLEKQ 66

RESULT 30
Q54660 PRELIMINARY; PRT; 79 AA.
ID   Q54660;
AC   Q54660;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Emml protein (Fragment).
GN   Name=emml;
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.

```

```

RC   STRAIN=M type 61;
RX   MEDLINE=95349390; PubMed=7623660;
RA   Whatmore A.M.; Kapur V.; Musser J.M.; Kehoe M.A.;
RT   "Molecular population genetic analysis of the emm subdivision of group
RT   A streptococcal emm-like genes: horizontal gene transfer and
RT   restricted variation among emm genes.";
RL   Mol. Microbiol. 15:1039-1048(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=M type 61;
RA   Whatmore A.M.;
RT   "Sequence Analysis of the Emm-Like Gene Family of Streptococcus
RT   pyogenes.";
RL   Thesis (1993), Microbiology, University of Newcastle Upon Tyne.
DR   EMBL; U20844; AAA87925.1; -.
FT   NON_TER 1
RT   NON_TER 79
SQ   SEQUENCE 79 AA; 9037 MW; 90DAAA62D93C505B CRC64;

Query Match
Best Local Similarity 42.0%; Score 37; DB 2; Length 79;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 QKAAELNQSKELEQQ 18
   :|:|:|:|:|:|
Db 44 KKLSEOEKSKNLEKQ 59

RESULT 31
Q90QH4 PRELIMINARY; PRT; 80 AA.
ID   Q90QH4;
AC   Q90QH4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Gag protein (Fragment).
GN   Name=gag;
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21248068; PubMed=11350662; DOI=10.1089/0892220151126526;
RA   Kaufmann G.R.; Suzuki K.; Cunningham P.; Mukai M.; Kondo M.;
RA   Imai M.; Zaunders J.; Cooper D.A.;
RT   "Impact of HIV type 1 protease, reverse transcriptase, cleavage site,
RT   and p6 mutations on the virological response to quadruple therapy with
RT   saquinavir, zidovudine, and two nucleoside analogs.";
RL   AIDS Res. Hum. Retroviruses 17:487-497(2001).
DR   EMBL; AF323215; AAK6678.1; -.
FT   NON_TER 1
RT   NON_TER 80
SQ   SEQUENCE 80 AA; 9086 MW; A8DF770A41C73090 CRC64;

Query Match
Best Local Similarity 42.0%; Score 37; DB 2; Length 80;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 PTKAAELNQSKELEQ 15
   :|:|:|:|:|:|
Db 49 P50KQEQIDEXQBL 63

RESULT 32
Q98PH3 PRELIMINARY; PRT; 81 AA.
ID   Q98PH3;
AC   Q98PH3;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein MYPV_7490.
GN   OrderedLocustNames=MYPV_7450;
OC   Mycoplasma pulmonis.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

```

```

AC 087UN3;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein VP0036.
GN OrderedLocustNames=VP0036;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIWD 020453 / Serotype O3:K6;
EX MEDLINE=22508445; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Nishida M., Nakano M., Yamashita A., Kubota Y., Kikuta S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae";
RL Lancet 361:743-749(2003).
DR EMBL; AP005073; BACS8299.1; -.
DR InterPro; IPR007475; DUF526.
DR Pfam; PF04380; DUF526; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 83 AA; 9631 MW; 9B1014610F579324 CRC64;

Query Match 42.0%; Score 37; DB 2; Length 83;
Best Local Similarity 43.8%; Pred.No. 7,4e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 3 OKAEINOKSKELFEQ 18
Db 64 OKLTEMERKUSELBEK 79

RESULT 35
Q54588 PRELIMINARY; PRT; 88 AA.
AC Q54588;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Emul protein (Fragment).
DE Name=emul;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M type 59;
RX MEDLINE=95198537; PubMed=7891551;
RA Whatmore A.M., Kapur V., Sullivan D.J., Musser J.M., Kehoe M.A.;
RT "Non-congruent relationships between variation in emm gene sequences
RT and the population genetic structure of group A streptococci.";
RL Mol. Microbiol. 14:619-631(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=M type 59;
RA Whatmore A.M.;
RT "Sequence Analysis of the Emm-Like Gene Family of Streptococcus
RT Pyogenes.";
RL Thesis (1993), Microbiology, University of Newcastle Upon Tyne.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=M type 59;
RA Whatmore A. M.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U11987; AAA99603.1; -.
DR PIR; S60837; S60837.1
FT NON_TER 1 88
FT TER 88
SQ SEQUENCE 88 AA; 10020 MW; C8B9B07BA5EE431E CRC64;

```

Query Match 42.0%; Score 37; DB 2; Length 88;
 Best Local Similarity 58.3%; Pred. No. 7.8e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 7 ELNKSKELEQQ 18
 ||:||||:
 Db 57 ELEKKNKELDSQ 68

RESULT 36
 Q54644 PRELIMINARY; PRT; 92 AA.
 ID Q54644
 AC Q54644;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE Emml protein (Fragment).
 GN Name=emml;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M type 4;
 RX MEDLINE=95349390; PubMed=7623660;
 RA Whitmore A.M., Kapur V., Musser J.M., Kehoe M.A.;
 RT "Molecular population genetic analysis of the emm subdivision of group
 RT A streptococcal emm-like genes: horizontal gene transfer and
 RT restricted variation among emm genes";
 RL Mol. Microbiol. 15:1039-1048(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M type 4;
 RA Whitmore A.M.;
 RT "Sequence Analysis of the Emm-like Gene Family of Streptococcus
 RT Pyogenes";
 RL Thesis (1993), Microbiology, University of Newcastle Upon Tyne.
 DR EMBL; U20828; AAA87909.1; -.
 FT NON_TER 1
 FT - NON_TER 92
 SQ SEQUENCE 92 AA; 10396 MW; BCEC2F44B1A032E7 CRC64;

Query Match 42.0%; Score 37; DB 2; Length 92;
 Best Local Similarity 50.0%; Pred. No. 8.2e+02;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 3 QKAELNOKSKLEQQ 18
 ||:||||:
 Db 54 KKEEBEKSKNLEKQ 69

RESULT 37
 Q9AOS8 PRELIMINARY; PRT; 92 AA.
 ID Q9AOS8
 AC Q9AOS8;
 DT 01-JUN-2001 (TRENBLREL. 17, Created)
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
 DE Muconolactone isomerase (EC 5.3.3.4).
 GN Name=cac3;
 OS Burkholderia sp. NK8.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia.
 OX NCBI_TaxID=140098;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NK8;
 RX MEDLINE=21097249; PubMed=11160806;
 RA Francisco P.B., Ogawa N., Suzuki K., Miyashita K.;
 RT "The chlorobenzoate dioxygenase genes of Burkholderia sp. strain NK8
 RT involved in the catabolism of chlorobenzoates";

RL Microbiology 147:121-133(2001).
 DR EMBL; AB024746; BAB21459.1; -.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 DR GO; GO:0016159; F:muconolactone delta-isomerase activity; IEA.
 DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
 DR InterPro; IPR011008; Dimer_A_B_barrel.
 DR InterPro; IPR003464; M1ase.
 DR Pfam; PF02426; M1ase; 1.
 DR Prodom; PD006619; M1ase; 1.
 KM Isomerase.
 SQ SEQUENCE 92 AA; 10764 MW; 267FD854AE81644 CRC64;

Query Match 42.0%; Score 37; DB 2; Length 92;
 Best Local Similarity 38.9%; Pred. No. 8.2e+02;
 Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 PTKAELNOKSKLEQQ 18
 ||:||||:
 Db 16 PAAKADEIKAREKESQE 33

RESULT 38
 Q8PT46 PRELIMINARY; PRT; 94 AA.
 ID Q8PT46
 AC Q8PT46;
 DT 01-OCT-2002 (TRENBLREL. 22, Created)
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
 DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)
 DE Methyltransferase.
 GN OrderedLocustNames=MM2872;
 OS Methanosarcina mazel (Methanosarcina frisia).
 OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2209;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Goe1 / Goe1 / ATCC BAA-199 / DSM 3647 / OCM 88;
 RX MEDLINE=22120827; PubMed=12125824;
 RA Dappenmeier U., Johann A., Hartisch T., Merkl R., Schmitz R.A.,
 RA Martinez-Arias R., Henne A., Wietzer A., Beumer S., Jacobl C.,
 RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Seckel S.,
 RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
 RA Fritz H.-J., Gottschalk G.;
 RT "The genome of Methanosarcina mazel: evidence for lateral gene
 RT transfer between Bacteria and Archaea";
 RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
 DR EMBL; AB013538; AAM32568.1; -.
 DR InterPro; IPR009078; Ferritin/RR_like.
 KM Complete proteome.
 SQ SEQUENCE 94 AA; 10708 MW; FF3D764953BF5107 CRC64;

Query Match 42.0%; Score 37; DB 2; Length 94;
 Best Local Similarity 43.8%; Pred. No. 8.4e+02;
 Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 3 QKAELNOKSKLEQQ 18
 ||:||||:
 Db 76 QOKOELBESKVEBEE 91

RESULT 39
 Q54540 PRELIMINARY; PRT; 94 AA.
 ID Q54540
 AC Q54540;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
 DE Emml protein (Fragment).
 GN Name=emml;
 OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;

KA Mehl N., Paine A., Berriman M., Churcher C., Harris B., Harris D., RA Mullin N., 2022/5/10/8; PMID=12358867; DOI=10.1093/nature/10935/

RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K., RA Buckee C.O., Burrows C., Cheruvach I., Chillingworth C., RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corcoran C., RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,

Search completed: August 29, 2005, 17:04:41
Job time : 177 secs

Search completed: August 29, 2005, 17:04:41
Job time : 177 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2005, 16:50:35 ; Search time 38 Seconds
(without alignments)
45.576 Million cell updates/sec

Title: US-10-031-289-1331
Perfect score: 88
Sequence: 1 PTQKAEINQSKSELEQQ 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 289416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 37678

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	48.9	74	2	SG0849
2	41	46.6	76	2	T13117
3	40	45.5	83	2	H82356
4	38	43.2	95	2	I53275
5	38	43.2	95	2	B25471
6	37	42.0	67	2	A95255
7	37	42.0	67	2	H98119
8	37	42.0	75	2	S61806
9	37	42.0	79	2	S61802
10	37	42.0	79	2	S62072
11	37	42.0	81	2	B90605
12	37	42.0	82	2	S61799
13	37	42.0	83	2	S61824
14	37	42.0	88	2	S60837
15	37	42.0	94	2	S60794
16	37	42.0	96	2	S60842
17	36	40.9	67	2	F83667
18	36	40.9	72	2	T26170
19	36	40.9	75	2	S60857
20	35	39.8	27	2	I54247
21	35	39.8	36	2	S07622
22	35	39.8	70	2	H72212
23	35	39.8	78	2	G82225
24	35	39.8	79	2	S61071
25	35	39.8	86	2	T43977
26	35	39.8	96	1	F69457
27	35	39.8	96	2	T23262
28	34	38.6	61	2	H83614
29	34	38.6	68	1	IABY3

30	34	38.6	71	2	S75075	hypothetical prote
31	34	38.6	77	2	S00970	keratin protein - pl
32	34	38.6	79	2	S78282	ribosomal protein
33	34	38.6	83	2	S62075	M-like protein enm
34	34	38.6	88	2	S60814	M protein precurs
35	34	38.6	89	2	S61821	M-like protein enm
36	34	38.6	94	2	S61805	M-like protein enm
37	34	38.6	96	2	S61827	M-like protein enm
38	34	38.6	96	2	S60811	M protein precurs
39	34	38.6	96	2	T18138	hypothetical prote
40	34	38.6	97	1	ISPSMP	muconolactone Delt
41	34	38.6	98	2	E97019	hypothetical prote
42	34	38.6	100	2	T10032	hypothetical prote
43	33.5	38.1	67	2	C71726	hypothetical prote
44	33	37.5	36	2	G81853	hypothetical prote
45	33	37.5	42	2	E82289	hypothetical prote
46	33	37.5	50	2	G89984	hypothetical prote
47	33	37.5	60	2	S60851	M protein precurs
48	33	37.5	66	1	RSB52L	ribosomal protein
49	33	37.5	66	2	S61083	M protein precurs
50	33	37.5	72	2	F83943	hypothetical prote
51	33	37.5	79	2	A84976	30S ribosomal prot
52	33	37.5	79	2	C70214	lipoprotein homolo
53	33	37.5	80	2	G64030	hypothetical prote
54	33	37.5	86	2	S62077	M-like protein enm
55	33	37.5	87	2	AE1228	B. subtilis YshA p
56	33	37.5	87	2	AG1581	hypothetical prote
57	33	37.5	88	2	A82938	hypothetical prote
58	33	37.5	90	2	I61689	myosin - human (fr
59	33	37.5	90	2	S62074	M-like protein enm
60	33	37.5	92	2	S22829	cOPR protein - Str
61	33	37.5	92	2	T31029	hypothetical prote
62	33	37.5	95	2	I67413	osteocalcin - mus
63	33	37.5	95	2	I61188	osteocalcin-relate
64	33	37.5	99	2	G82269	conserved hypotet
65	33	37.5	100	2	C89989	hypothetical prote
66	32.5	36.9	44	2	E89877	hypothetical prote
67	32.5	36.9	51	2	C97077	hypothetical prote
68	32.5	36.9	51	2	S61847	hnpX protein - Pse
69	32	36.4	34	2	C97368	hypothetical prote
70	32	36.4	38	2	S78339	photosystem II pro
71	32	36.4	40	2	S58853	homeotic protein u
72	32	36.4	43	2	D97484	hypothetical prote
73	32	36.4	44	2	B64576	hypothetical prote
74	32	36.4	46	2	E69830	hypothetical prote
75	32	36.4	49	1	A61280	osteocalcin - rabb
76	32	36.4	45	1	E69970	spore coat protein
77	32	36.4	66	1	RSB529	ribosomal protein
78	32	36.4	67	2	I54386	merlin protein - r
79	32	36.4	74	2	S60836	M protein precurs
80	32	36.4	84	2	AG0385	exodeoxyribonuclea
81	32	36.4	88	2	S60812	M protein precurs
82	32	36.4	95	2	S61076	M protein precurs
83	32	36.4	95	2	H83237	hypothetical prote
84	32	36.4	96	2	I17621	Ig heavy chain V r
85	31.5	35.8	66	2	G71834	ribosomal protein
86	31.5	35.8	66	2	G64683	ribosomal protein
87	31	35.2	20	2	JP0070	ribosomal protein
88	31	35.2	22	2	JP0071	ribosomal protein
89	31	35.2	44	2	T26893	hypothetical prote
90	31	35.2	50	2	S72456	pachytene arrest p
91	31	35.2	53	2	A90599	hypothetical prote
92	31	35.2	56	2	T40042	truncated endonuc
93	31	35.2	57	2	S19088	dihydropyrimidin d
94	31	35.2	59	2	S78730	protein YML081c-a
95	31	35.2	61	2	B81860	hypothetical prote
96	31	35.2	62	2	I50495	light meromyosin -
97	31	35.2	63	1	RSBC29	ribosomal protein
98	31	35.2	63	2	A11006	50S ribosomal chai
99	31	35.2	63	2	A91151	50S ribosomal subu
100	31	35.2	63	2	E85996	50S ribosomal subu

ALIGNMENTS

RESULT 1

S60849 M protein precursor - Streptococcus pyogenes (serotype M65) (fragment)

C:Species: Streptococcus pyogenes

A:Variety: serotype M65

C>Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C/Accession: S60849

R:Watmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.

Mol. Microbiol. 14, 619-631, 1994

A>Title: Non-congruent relationships between variation in emm gene sequences and the pop

A:Reference number: S60784; MUID:95198537; PMID:7891551

A:Accession: S60849

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-74 <MHA>

A:Cross-references: UNIPROT:Q54581; EMBL:U11980

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994

C:Superfamily: M5 protein

Query Match 48.9%; Score 43; DB 2; Length 74;
Best Local Similarity 47.1%; Pred. No. 15;

Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 TOKAEINOKSKELEQ 18

DB 56 TEKNEELDKKKKJDSQ 72

RESULT 2

T13117 protein gp31 - phage N15

C:Species: phage N15

C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004

C/Accession: T13117

R:Hendrix, R.W.; Ravin, V.K.; Casjens, S.R.; Ford, M.E.; Ravin, N.V.; Smirnov, I.K.

submitted to the EMBL Data Library, May 1998

A:Reference number: Z17603

A:Accession: T13117

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-76 <HEN>

A:Cross-references: UNIPROT:O64342; EMBL:AF064539; NID:g3192683; PID:g3192720; PIDN:AACT

C:Genetics:

A>Note: gene 31

Query Match 46.6%; Score 41; DB 2; Length 76;
Best Local Similarity 56.2%; Pred. No. 31;

Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 OKAEINOKSKELEQ 18

DB 43 QEAAILQVARELTQE 58

RESULT 3

H82356 conserved hypothetical protein VC0163 (imported) - Vibrio cholerae (strain N16961 serogr

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C/Accession: H82356

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: H82356

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-83 <HEI>

A:Cross-references: UNIPROT:Q9KV13; GB:AE004106; GB:AE003852; NID:9654561; PIDN:AE9333;
A:Experimental source: serogroup O1, strain N16961; biotype El Tor
C:Genetics:

A:Gene: VC0163

A:Map position: 1

C:Superfamily: Escherichia coli hypothetical protein b3042

Query Match 45.5%; Score 40; DB 2; Length 83;
Best Local Similarity 64.3%; Pred. No. 47;

Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 OKAEINOKSKELE 16

DB 64 QKTELEQKAELE 77

RESULT 4

I53275 osteocalcin - mouse

C:Species: Mus sp. (mouse)

C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 20-Aug-1999

C/Accession: I53275

R:Rahman, S.; Oberdorf, A.; Montecino, M.; Tanhauser, S.M.; Lian, J.B.; Stein, G.S.; Lai

Endocrinology 133, 3050-3053, 1993

A>Title: Multiple copies of the bone-specific osteocalcin gene in mouse and rat.

A:Reference number: I53275; MUID:94062692; PMID:8243336

A:Accession: I53275

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-95 <RES>

A:Cross-references: GB:S67455; NID:g456854; PIDN:AA829145.1; PID:g456856

C:Genetics:

A:Insertions: 22/1; 33/1; 52/2; 72/2

C:Superfamily: osteocalcin

Query Match 43.2%; Score 38; DB 2; Length 95;
Best Local Similarity 44.4%; Pred. No. 11e+02;

Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 PROKAEINOKSKELEQ 18

DB 63 PTRYQCELNPACDELSQ 80

RESULT 5

B25471 osteocalcin precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004

C/Accession: B25471; A49871; I61189

R:Celeste, A.J.; Rosen, V.; Becker, J.L.; Kriz, R.; Wang, E.A.; Wozney, J.M.

EMBO J. 5, 1885-1890, 1986

A>Title: Isolation of the human gene for bone gla protein utilizing mouse and rat cDNA c]

A:Reference number: A91045; MUID:87004555; PMID:3019668

A:Accession: B25471

A:Molecule type: DNA

A:Residues: 1-95 <CEL>

A:Cross-references: UNIPROT:P04641

R:Desbois, C.; Hogue, D.A.; Karsenty, G.

J. Biol. Chem. 269, 1183-1190, 1994

A>Title: The mouse osteocalcin gene cluster contains three genes with two separate spaci

A:Reference number: A49871; MUID:94117426; PMID:8288580

A:Accession: A49871

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-95 <RES>

A:Cross-references: GB:L24429; NID:g455452; PIDN:AAA39854.1; PID:g455453

A:Accession: I61189

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3, '1F', 6-10, 'T', 12-95 <RE2>

A:Cross-references: GB:L24431; NID:g455456; PIDN:AAA39856.1; PID:g455457

C:Genetics:


```

07 0 AELNQNKNLEEDQ 18
Db  ||: ||: ||: ||:
38 AEIAQKREIEQR 50

```

C;Date: 09-Mar-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004

C;Accession: S62072; S61808

```

R:Whitmore, A.M. EMBL Data Library, February 1995
submitted to the EMBL Data Library, February 1995
A:Reference number: S62072
A:Accession: S62072
A:Molecule type: DNA
A:Residues: 1-79 <WHA>
A:Cross-references: UNIPROT:Q54642; EMBL:U20826; NID:9687748; PIDN:AAA87907.1; PID:9687748
A:Experimental source: serotype M44
R:Whitmore, A.M.; Kapur, V.; Musser, J.M.; Kehoe, M.A.
Mol. Microbiol. 15, 1039-1048, 1995
A:Title: Molecular population genetic analysis of the emn subdivision of group A streptococci
A:Reference number: S61799; MUID:95349390, PMID:7623660
A:Accession: S61808
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 9-57 <WHM>
A:Cross-references: EMBL:U20826
A:Experimental source: serotype M44
C:Genetic8:
A:Gene: emm44
C:Superfamily: M5 protein
C:Keywords: cell wall, virulence factor
F:1-21/Domains: signal sequence (fragment) #status predicted <SIG>
F:22-79/Product: protein M-like protein (fragment) #status predicted <WAT>

Query Match      42.0%; Score 37; DB 2; Length 79;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Cy 3 OKAAELNOKSKLEEQ 18
   :|::|||::|:
Db 55 KLEEBEOKSKNLEKQ 70

RESULT 11
E90605
hypothetical protein MYPV_7490 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: E90605
A:Cross-references: UNIPROT:Q98PH3; GB:AL445566; PID:g14090164; PIDN:CAC13922.1; GSPDB:G
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis
A:Reference number: A9512; MUID:21267165; PMID:11353084
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis
A:Reference number: A9512; MUID:21267165; PMID:11353084
A:Status: preliminary
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-81 <KUR>
A:Cross-references: UNIPROT:Q98PH3; GB:AL445566; PID:g14090164; PIDN:CAC13922.1; GSPDB:G
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPV_7490
A:Genetic code: SGC3

Query Match      42.0%; Score 37; DB 2; Length 81;
Best Local Similarity 53.3%; Pred. No. 1.3e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Cy 2 TOKAAELNOKSKLE 16
   ||::|||
Db 49 TLKMKVNOHKKELE 63

RESULT 12
S61799
M-like protein emn precursor - Streptococcus pyogenes (serotype M61) (fragment)
C:Species: Streptococcus pyogenes
A:Variety: serotype M61
C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jul-2004
C:Accession: S61799; S62079
R:Whitmore, A.M.; Kapur, V.; Musser, J.M.; Kehoe, M.A.
Mol. Microbiol. 15, 1039-1048, 1995
A:Title: Molecular population genetic analysis of the emn subdivision of group A streptococci

```

```

A.Reference number: S61799; MUID:95349390; PMID:7623660
A.Accession: S61799
A.Status: not compared with conceptual translation
A.Molecule type: DNA
A.Residues: 1-49 <WHA>
A.Cross-references: UNIPROT:O53476
A.Experimental source: strain M61-PHLIS; serotype M61
A.Note: Only a part of the coding sequence is given
R.Whitmore, A.M.
submitted to the EMBL Data Library, February 1995
A.Reference number: S62072
A.Accession: S62072
A.Molecule type: DNA
A.Residues: 3-82 <WHA2>
A.Cross-references: EMBL:U20844
A.Experimental source: strain M61-PHLIS; serotype M61
C.Genetics:
A.Gene: emn
C.Superfamily: M5 protein
C.Keywords: cell wall; surface antigen; virulence
F.1-13/Domant: signal sequence (fragment) #status predicted <SIG>
F.14-82/Product: M protein-like sequence (fragment) #status predicted <MAT>

Query Match      42.0%; Score 37; DB 2; Length 82;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches      8; Conservative      3; Mismatches      5; Indels      0; Gaps      0;

OY      3 OKAAELNOKSKLEEQO 18
       :|::|||::|:
Db      47 KKLSEEGEKSKYLEKQ 62

RESULT 13
S61824
M.-like protein emn precursor - Streptococcus pyogenes (serotype MPR4245) (fragment)
C.Species: Streptococcus pyogenes
A.Variety: serotype MPR4245
C.Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C.Accession: S61824
R.Whitmore, A.M.; Kapur, V.; Musser, J.M.; Kehoe, M.A.
MOL. Microbiol. 15, 1039-1048, 1995
A.Title: Molecular population genetic analysis of the emn subdivision of group A streptococci
A.Reference number: S61799; MUID:95349390; PMID:7623660
A.Accession: S61824
A>Status: nucleic acid sequence not shown; translation not shown
A.Molecule type: DNA
A.Residues: 1-83 <WHA>
A.Cross-references: UNIPROT:O54643; EMBL:U20827; NID:G687750; PIDN:AAA87908.1; PID:G68775
A.Experimental source: strain R90/3355; serotype MPR4245
A.Note: The nucleotide sequence was submitted to the EMBL Data Library, February 1995
C.Genetics:
A.Gene: emn
C.Superfamily: M5 protein
C.Keywords: cell wall; surface antigen; virulence

Query Match      42.0%; Score 37; DB 2; Length 83;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches      8; Conservative      3; Mismatches      5; Indels      0; Gaps      0;

OY      3 OKAAELNOKSKLEEQO 18
       :|::|||::|:
Db      53 KKLSEEGEKSKYLEKQ 68

RESULT 14
S60837
M protein precursor - Streptococcus pyogenes (serotype M59) (fragment)
C.Species: Streptococcus pyogenes
A.Variety: serotype M59
C.Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C.Accession: S60837
R.Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.
MOL. Microbiol. 14, 619-631, 1994

```

A:Title: Non-congruent relationships between variation in emm gene sequences and the pop
 A:Reference number: S60784; MUID:95198537; PMID:7891551
 A:Accession: S60837
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-88 <MHA>
 A:Cross-references: UNIPROT:O54588; EMBL:U11987; NID:G533651; PIDN:AAA99603.1; PID:q1235
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
 C:Superfamily: M5 protein

Query Match 42.0%; Score 37; DB 2; Length 88;
 Best Local Similarity 58.3%; Pred. No. 1.4e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 7 ELNOKSKELBOQ 18
 ||:|:|:|:
 Db 57 ELEKKNKELDSQ 68

RESULT 15

S60794
 M protein precursor - Streptococcus pyogenes (serotype M13) (fragment)

C:Species: Streptococcus pyogenes

A:Variety: serotype M13

C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C:Accession: S60794

R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.

Mol. Microbiol. 14, 619-631, 1994

A:Title: Non-congruent relationships between variation in emm gene sequences and the pop

A:Reference number: S60784; MUID:95198537; PMID:7891551

A:Accession: S60794

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-94 <MHA>

A:Cross-references: UNIPROT:O54540; EMBL:U11936; NID:G533549; PIDN:AAA99552.1; PID:q1235

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994

C:Superfamily: M5 protein

Query Match 42.0%; Score 37; DB 2; Length 94;
 Best Local Similarity 58.3%; Pred. No. 1.5e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 7 ELNOKSKELBOQ 18
 ||:|:|:|:
 Db 68 ELEKKNKELDSQ 79

RESULT 16

S60842
 M protein precursor - Streptococcus pyogenes (serotype M63) (fragment)

C:Species: Streptococcus pyogenes

A:Variety: serotype M63

C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C:Accession: S60842

R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.

Mol. Microbiol. 14, 619-631, 1994

A:Title: Non-congruent relationships between variation in emm gene sequences and the pop

A:Reference number: S60784; MUID:95198537; PMID:7891551

A:Accession: S60842

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-86 <MHA>

A:Cross-references: UNIPROT:O54583; EMBL:U11982; NID:G533641; PIDN:AAA99598.1; PID:q1235

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994

C:Superfamily: M5 protein

Query Match 42.0%; Score 37; DB 2; Length 96;
 Best Local Similarity 58.3%; Pred. No. 1.5e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 7 ELNOKSKELBOQ 18
 ||:|:|:|:
 Db 70 ELEKKNKELDSQ 81

RESULT 17

F83667
 ribosomal protein L29 rpmC [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C:Accession: F83667

R:Takam, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Saeki, R.; Masui, N.; Fuji, F.; Hiran

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: F83667

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-67 <STO>

A:Cross-references: UNIPROT:Q929K6; GB:AF001507; GB:BA000004; NID:q10172612; PIDN:BA0386

A:Experimental source: strain C-125

C:Genetics:

A:Gene: rpmC

C:Superfamily: Escherichia coli ribosomal protein L29

Query Match 40.9%; Score 36; DB 2; Length 67;
 Best Local Similarity 46.2%; Pred. No. 1.5e+02;
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Oy 6 AEIENOKSKELBOQ 18
 ||:|:|:|:
 Db 12 AEIEQKTKSLKEE 24

RESULT 18

T26170
 hypothetical protein W04G5.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T26170

R:Kershaw, J.

submitted to the EMBL Data Library, March 1997

A:Reference number: Z20164

A:Accession: T26170

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-72 <MIL>

A:Cross-references: UNIPROT:O18168; EMBL:Z93391; PIDN:CA807680.1; GSPDB:GN00019; CESP:W04

A:Experimental source: clone W04G5

C:Genetics:

A:Gene: CESP:W04G5.3

A:Map position: 1

A:Introns: 20/1

Query Match 40.9%; Score 36; DB 2; Length 72;
 Best Local Similarity 53.8%; Pred. No. 1.6e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 1 PTOKAELNOKSK 13
 ||:|:|:|:
 Db 21 PSKLEELQKKK 33

RESULT 19

S60857
 M protein precursor - Streptococcus pyogenes (serotype TR2612) (fragment)

C:Species: Streptococcus pyogenes

A:Variety: serotype TR2612

C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C:Accession: S60857

R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.

Mol. Microbiol. 14, 619-631, 1994

A:Title: Non-congruent relationships between variation in emm gene sequences and the pop

A:Reference number: S60784; MUID:95198537; PMID:7891551

A:Accession: S60857

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-75 <MHA>
A:Cross-references: UNIPROT:Q54553; EMBL:U11950; NID:G533577; PIDN:AAA9566.1; PID:91235
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Superfamily: M5 protein

Query Match 40.9%; Score 36; DB 2; Length 75;
Best Local Similarity 46.7%; Pred. No. 1.7e+02;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 4 KAAELNOKSKLEEQ 18
|||::|||::|
Db 61 KNKELDKKKKELDSR 75

RESULT 20
154247
GTP-binding regulatory protein Gt alpha-2 chain - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C:Accession: I54247
R:Hitano, T.
Hokkaido Igaku Zasshi 68, 885-893, 1993
A:Title: [Analysis of cell specific transcription of the human cone transducin alpha sub
A:Reference number: I54247; MUID:94156312; PMID:8112713
A:Accession: I54247
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-27 <RES>
A:Cross-references: UNIPROT:Q16162; GB:S70567; NID:G545549; PIDN:AMD14063.1; PID:9426176
C:Superfamily: GTP-binding regulatory protein Ge alpha chain

Query Match 39.8%; Score 35; DB 2; Length 27;
Best Local Similarity 58.3%; Pred. No. 87;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 7 ELNOKSKLEEQ 18
|||::|||::|
Db 11 ELAKRSKLEKK 22

RESULT 21
S07622
avenin gamma-4 - small naked oat (fragment)
N:Alternate names: prolamin, 25K; seed storage protein, 25K
C:Species: Avena nuda (small naked oat)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 12-Apr-1995
C:Accession: S07622
R:Pernollet, J.C.; Huec, J.C.; Galle, A.M.; Sallantin, M.
Biochimie 69, 683-689, 1987
A:Title: N-terminal sequences of oat avenins compared to other cereal prolamins.
A:Reference number: S06211; MUID:88078111; PMID:3120802
A:Accession: S07622
A:Molecule type: protein
A:Residues: 1-36 <PER>
C:Superfamily: gliadin
C:Keywords: prolamin; seed; storage protein

Query Match 39.8%; Score 35; DB 2; Length 36;
Best Local Similarity 44.4%; Pred. No. 1.1e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 PTOKAAELNOKSKLEEQ 18
|||::|||::|
Db 9 PEQQQPFLLQQPFLQLQQ 26

RESULT 22
H72212
exodeoxyribonuclease, small subunit - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: H72212

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: H72212
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-70 <ARN>
A:Cross-references: GB:AE001815; GB:AE000512; NID:G4982341; PIDN:AMD36843.1; PID:G498235
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1769

Query Match 39.8%; Score 35; DB 2; Length 70;
Best Local Similarity 43.8%; Pred. No. 2.2e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 3 OKAAELNOKSKLEEQ 18
::|||::|||::|
Db 27 ERGVELYRKCKELTQQ 42

RESULT 23
G82225
hypothetical protein VC1223 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82225
R:Haideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Baas, S.; Qin, H.; Dragol, I.; Sellers, P.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82225
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-78 <HET>
A:Cross-references: UNIPROT:Q9KSN3; GB:AE004202; GB:AE003852; NID:G9655698; PIDN:AAE9438;
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1223
A:Map position: 1

Query Match 39.8%; Score 35; DB 2; Length 78;
Best Local Similarity 37.5%; Pred. No. 2.4e+02;
Matches 6; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 PTOKAAELNOKSKLE 16
|||::|||::|
Db 33 PTOAASSTWKARSLD 48

RESULT 24
S61071
M protein precursor - Streptococcus pyogenes (serotype M81) (fragment)
C:Species: Streptococcus pyogenes
A:Variety: serotype M81
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S61071
R:Wattmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.
Mol. Microbiol. 14, 619-631, 1994
A:Title: Non-congruent relationships between variation in emm gene sequences and the pop
A:Reference number: S60784; MUID:95198537; PMID:7891551
A:Accession: S61071
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-79 <MHA>
A:Cross-references: UNIPROT:Q54604; EMBL:U12003; NID:G533683; PIDN:AAA99619.1; PID:91235f
A:Note: the nucleotide sequence was submitted to the EMBL Data Library July 1994
C:Superfamily: M5 protein

Query Match 39.8%; Score 35; DB 2; Length 79;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 7 ELNOKSKELE 16
|||:|||||
Db 57 ELNENKRLTE 66

RESULT 25
143977
hypothetical protein U17 [imported] - human herpesvirus 6 (strain HST)
C:Species: human herpesvirus 6
A:Variety: strain HST
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43977
R:Jesgawa, Y.; Mukai, T.; Nakano, K.; Kagawa, M.; Chen, J.; Mori, Y.; Sunagawa, T.; Kawa
J. Virol. 73, 8053-8063, 1999
A:Title: Comparison of the complete DNA sequences of human herpesvirus 6 variants A and
A:Reference number: 222732; MUID:99412319; PMID:10482554
A:Accession: T43977
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-86 <ISE>
A:Cross-references: UNIPROT:Q9WT47; EMBL:AB021506; NID:G4395977; PIDN:BA478238.1; PID:94
A:Experimental source: strain HST; pop. variant B
C:Genetics:
A>Note: U17

Query Match 39.8%; Score 35; DB 2; Length 86;
Best Local Similarity 53.8%; Pred. No. 2.7e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PTKAABLQEMCK 13
|||:|||||
Db 66 PTKSAAELQEMCK 78

RESULT 26
F69457
conserved hypothetical protein AF1663 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: F69457
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Globeck, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uitterlbeck, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Moese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: F69457
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-96 <KLE>
A:Cross-references: UNIPROT:O28610; GB:AE000989; GB:AE000782; NID:G2689312; PIDN:AAB8958
C:Superfamily: hypothetical protein AF2083

Query Match 39.8%; Score 35; DB 1; Length 96;
Best Local Similarity 43.8%; Pred. No. 3e+02;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 TOKAABLQKSELEQ 17
|||:|||||
Db 67 TEKGRFPIQKSEIEB 82

RESULT 27
T23262
hypothetical protein K0388.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23262
R:White, S.
submitted to the EMBL Data Library, June 1996
A:Reference number: 219718
A:Accession: T23262
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-96 <WIL>
A:Cross-references: UNIPROT:Q21177; EMBL:Z74039; PIDN:CAA98500.1; GSPDB:GN00023; CESP:KO
A:Experimental source: clone K0388
C:Genetics:
A:Gene: CESP:K0388.4
A:Map position: 5
A:Introns: 37/2

Query Match 39.8%; Score 35; DB 2; Length 96;
Best Local Similarity 63.6%; Pred. No. 3e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 7 ELNOKSKELEQ 17
|||:|||||
Db 42 KVNQKRELEQ 52

RESULT 28
H83614
hypothetical protein PA0258 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: H83614
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bru
adman, S.; Yun, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidig, K.; Lim,
L.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathoge
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H83614
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-61 <STO>
A:Cross-references: UNIPROT:Q916N1; GB:AB004463; GB:AE004091; NID:G9946086; PIDN:AG0364
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0258

Query Match 38.6%; Score 34; DB 2; Length 61;
Best Local Similarity 77.8%; Pred. No. 2.7e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 10 QKSKLEEQ 18
|||:|||||
Db 11 QRIKLEEQ 19

RESULT 29
IABY3
Proteinase A inhibitor 3 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YM8010.04; protein YM8174c
C:Species: Saccharomyces cerevisiae
C>Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 09-Jul-2004
C:Accession: A01334; S16697; S55121
R:Biedeman, K.; Montali, U.; Martin, B.; Svendsen, I.; Ottesen, M.
Carlsberg Res. Commun. 45, 225-235, 1980
A:Title: The amino acid sequence of proteinase A inhibitor 3 from baker's yeast.
A:Reference number: A01334
A:Accession: A01334
A:Molecule type: protein
A:Residues: 1-68 <BIE>
A:Cross-references: UNIPROT:P01094
A>Note: nearly all of the inhibitory activity is present in the peptide consisting of res
R:Schu, P.; Wolf, D.H.
FEBS Lett. 283, 78-84, 1991

A/Title: The proteinase YscA-inhibitor, I(A) (3), gene. Studies of cytoplasmic proteinase
 A/Reference number: S16692; MUID:91243884; PMID:2037077
 A/Accession: S16692
 A/Molecule type: DNA
 A/Residues: 1-68 <SCH>
 A/Cross-references: EMBL:X60050; NID:g4094; PIDN:CAA42650.1; PID:g4095
 R/Churher, C.M.
 Submitted to the EMBL Data Library, June 1995
 A/Reference number: S55118
 A/Accession: S55121
 A/Molecule type: DNA
 A/Residues: 1-68 <CHU>
 A/Cross-references: EMBL:Z49808; NID:g854440; PIDN:CAA89907.1; PID:g854444; GSPDB:GN0001
 A/Experimental source: strain AB972
 C/Genetics:
 A/Gene: SGD:PA13; MIPS:YKR174C
 A/Cross-references: SGD:S0004786; MIPS:YKR174C
 A/Map position: 13R
 C/Superfamily: proteinase A inhibitor 3
 C/Keywords: acetylated amino end; blocked amino end; proteinase inhibitor
 F/1/Modified site: acetylated amino end (Met) #status experimental

Query Match 38.6%; Score 34; DB 1; Length 68;
 Best Local Similarity 58.3%; Pred. No. 3e+02;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 OKAAELNOKSKELQ 14
 DB 6 OKVSEIFOSKE 17

RESULT 30
 S75075
 hypochlorous protein sal0467 - *Synechocystis* sp. (strain PCC 6803)
 C/Species: *Synechocystis* sp.
 A/Variety: PCC 6803
 C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

A/Accession: S75075
 A/Molecule type: DNA
 A/Residues: 1-71 <KAN>
 A/Cross-references: UNIPROT:P73875; EMBL:D90910; GB:AB001339; NID:g1652956; PIDN:BA1793
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 38.6%; Score 34; DB 2; Length 71;
 Best Local Similarity 53.3%; Pred. No. 3.1e+02;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 4 KAAELNOKSKELQ 18
 DB 6 KARELLAKERONEQ 20

RESULT 31
 S00970
 kerA1 protein - plasmid RK2
 C/Species: plasmid RK2
 C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 08-Oct-1999
 A/Accession: S00970; B53306; D33387
 R/Thomas, C.M.; Ibbotson, J.P.; Wang, N.; Smith, C.A.; Tipping, R.; Loader, N.M.
 Nucleic Acids Res. 16, 5345-5359, 1988
 A/Title: Gene regulation on broad host range plasmid RK2: identification of three novel
 A/Reference number: S00970; MUID:86262548; PMID:283814
 A/Accession: S00970
 A/Molecule type: DNA
 A/Residues: 1-77 <THO>

A/Cross-references: EMBL:X07248; NID:g41866; PIDN:CAA30233.1; PID:g41867
 R/Kornacki, J.A.; Chang, C.H.; Figureki, D.H.
 J. Bacteriol. 175, 5078-5090, 1993
 A/Title: Kil-kor regulon of promiscuous plasmid RK2: structure, products, and regulation
 A/Reference number: A53306; MUID:93352413; PMID:8349548
 A/Accession: B53306
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-77 <KOR>
 A/Cross-references: GB:U18919; NID:g400385; PIDN:AAA92765.1; PID:g400387
 R/Kornacki, J.A.; Burlage, R.S.; Figureki, D.H.
 J. Bacteriol. 172, 3040-3050, 1990
 A/Title: The kil-kor regulon of broad-host-range plasmid RK2: nucleotide sequence, poly
 A/Reference number: A53387; MUID:90264294; PMID:2160936
 A/Accession: D35387
 A/Molecule type: DNA
 A/Residues: 1-16 <KO2>
 A/Cross-references: GB:M32794
 C/Genetics:
 A/Gene: kerA1; k1eA
 A/Genome: plasmid

Query Match 38.6%; Score 34; DB 2; Length 77;
 Best Local Similarity 37.5%; Pred. No. 3.4e+02;
 Matches 6; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

OY 3 OKAAELNOKSKELQ 18
 DB 32 EOPAEITROAEIDRHK 47

RESULT 32
 S78282
 ribosomal protein S16, chloroplast - *Odontella sinensis* chloroplast
 C/Species: chloroplast *Odontella sinensis*
 C/Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 09-Jul-2004

A/Accession: S78282
 A/Molecule type: DNA
 A/Residues: 1-79 <KOW>
 A/Cross-references: UNIPROT:P49503; EMBL:Z67753; NID:g1185127; PIDN:CAA91655.1; PID:g1185
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
 C/Genetics:
 A/Gene: rps16
 A/Genome: chloroplast
 C/Superfamily: *Escherichia coli* ribosomal protein S16
 C/Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 38.6%; Score 34; DB 2; Length 79;
 Best Local Similarity 41.2%; Pred. No. 3.5e+02;
 Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 1 PTQAAELNOKSKELQ 17
 DB 63 PTQVNLTKKAKTIDQ 79

RESULT 33
 S62075
 W-like protein em precursor - *Streptococcus pyogenes* (serotype M58) (fragment)
 C/Species: *Streptococcus pyogenes*
 A/Variety: serotype M58
 C/Date: 09-Mar-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 A/Accession: S62075; S61812
 R/Whitmore, A.M.
 submitted to the EMBL Data Library, February 1995
 A/Reference number: S62072
 A/Accession: S62075

A:Molecule type: DNA
A:Residues: 1-83 <WHA>
A:Cross-references: UNIPROT:Q54562; EMBL:U20836; NID:9694065; P1DN:AAA87917.1; PID:96940
A:Experimental source: strain M58-PHLS; serotype M58
R:Whitmore, A.M.; Kapur, V.; Musser, J.M.; Kehoe, M.A.
Mol. Microbiol. 15, 1039-1048, 1995
A:Title: Molecular population genetic analysis of the emm subdivision of group A streptococci
A:Reference number: S61799; MUID:95349390; PMID:7623660
A:Accession: S61812
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 7-48 <WH2>
A:Cross-references: EMBL:U20836
A:Experimental source: strain M58-PHLS; serotype M58
C:Genetics:
A:Gene: emm
C:Superfamily: M5 protein
C:Keywords: cell wall; surface antigen; virulence
F:1-19/Domain: signal sequence (fragment) #status predicted <SIG>
F:20-83/Product: M protein-like protein emm (fragment) #status predicted <MAT>
Query Match 38.6%; Score 34; DB 2; Length 83;
Best Local Similarity 43.8%; Pred. No. 3.6e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
Oy 3 OKAAELNOKSKLEEQ 18
Db 53 KLEEEQKSKNLEKE 68
RESULT 34
S60814
M protein precursor - Streptococcus pyogenes (serotype M41) (fragment)
C:Species: Streptococcus pyogenes
A:Variety: serotype M41
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S60814
R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.
Mol. Microbiol. 14, 619-631, 1994
A:Title: Non-congruent relationships between variation in emm gene sequences and the por
A:Reference number: S60784; MUID:95198537; PMID:7891551
A:Accession: S60814
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-88 <WHA>
A:Cross-references: UNIPROT:Q54569; EMBL:U11967; NID:9533611; P1DN:AAA9583.1; PID:91235
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Superfamily: M5 protein
Query Match 38.6%; Score 34; DB 2; Length 88;
Best Local Similarity 58.3%; Pred. No. 3.8e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Oy 7 ELNOKSKLEEQ 18
Db 66 EAEQKSKAEQK 77
RESULT 35
S61821
M-like protein emm precursor - Streptococcus pyogenes (serotype M76) (fragment)
C:Species: Streptococcus pyogenes
A:Variety: serotype M76
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S61821
R:Whitmore, A.M.; Kapur, V.; Musser, J.M.; Kehoe, M.A.
Mol. Microbiol. 15, 1039-1048, 1995
A:Title: Molecular population genetic analysis of the emm subdivision of group A streptococci
A:Reference number: S61799; MUID:95349390; PMID:7623660
A:Accession: S61821
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-89 <WHA>

A:Cross-references: UNIPROT:Q54668; EMBL:U20852; NID:9694097; P1DN:AAA87933.1; PID:96940
A:Experimental source: strain M76-PHLS; serotype M76
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995
C:Genetics:
A:Gene: emm
C:Superfamily: M5 protein
C:Keywords: cell wall; surface antigen; virulence
Query Match 38.6%; Score 34; DB 2; Length 89;
Best Local Similarity 43.8%; Pred. No. 3.9e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
Oy 3 OKAAELNOKSKLEEQ 18
Db 55 KLEEEQKSKNLEKE 70
RESULT 36
S61805
M-like protein emm precursor - Streptococcus pyogenes (serotype M27) (fragment)
C:Species: Streptococcus pyogenes
A:Variety: serotype M27
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S61805
R:Whitmore, A.M.; Kapur, V.; Musser, J.M.; Kehoe, M.A.
Mol. Microbiol. 15, 1039-1048, 1995
A:Title: Molecular population genetic analysis of the emm subdivision of group A streptococci
A:Reference number: S61799; MUID:95349390; PMID:7623660
A:Accession: S61805
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-94 <WHA>
A:Cross-references: UNIPROT:Q54666; EMBL:U20850; NID:9694093; P1DN:AAA87931.1; PID:96940
A:Experimental source: NCTC 8328; serotype M27
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995
C:Genetics:
A:Gene: emm
C:Superfamily: M5 protein
C:Keywords: cell wall; surface antigen; virulence
Query Match 38.6%; Score 34; DB 2; Length 94;
Best Local Similarity 43.8%; Pred. No. 4.1e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
Oy 3 OKAAELNOKSKLEEQ 18
Db 53 KLEEEQKSKNLEKE 68
RESULT 37
S61827
M-like protein emm precursor - Streptococcus pyogenes (serotype MPT3875) (fragment)
C:Species: Streptococcus pyogenes
A:Variety: serotype MPT3875
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S61827
R:Whitmore, A.M.; Kapur, V.; Musser, J.M.; Kehoe, M.A.
Mol. Microbiol. 15, 1039-1048, 1995
A:Title: Molecular population genetic analysis of the emm subdivision of group A streptococci
A:Reference number: S61799; MUID:95349390; PMID:7623660
A:Accession: S61827
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-96 <WHA>
A:Cross-references: UNIPROT:Q54645; EMBL:U20829; NID:9687754; P1DN:AAA87910.1; PID:96877
A:Experimental source: strain 3875-PHLS; serotype MPT3875
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995
C:Genetics:
A:Gene: emm
C:Superfamily: M5 protein
C:Keywords: cell wall; surface antigen; virulence
Query Match 38.6%; Score 34; DB 2; Length 96;

A;Status: preliminary;
A;Molecule type: DNA


```
A:Cross-references: UNIPROT:O33084; EMBL:I14967; NID:g3370268; PIDN:CAA75210.1; PII:g233  
Query Match      38.6%; Score 34; DB 2; Length 100;  
Best Local Similarity 41.2%; Pred. No. 4.3e+02;  
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
  
QY          2 TTKAAELNQSKKELEQQ 18  
             ||::||::|:||:  
Db         12 TTQAAGPDQIASGLSQE 28  
  
RESULT_43  
C71726  
hypothetical protein RP158 - Rickettsia prowazekii  
CISpecies: Rickettsia prowazekii  
CIDate: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004  
CAccession: C71726  
R.Anderson, S.G.E.; Zomorroddipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alemark, U.  
Nature 396, 133-140, 1998  
A>Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
A.Reference number: AVI630; MUID:99039499; PMID:9823893  
A.Accession: C71726  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A.Molecule type: DNA  
A.Residues: 1-67 <AND>  
A.CROSS-references: UNIPROT:O9ZE04; GB:AJ235270; GB:AJ235269; NID:g3860572; PIDN:CAA1462Z  
A.Experimental source: strain Madrid E  
A.Genetics:  
A.Gene: RP158
```

```

RESULT 44
GB1853
hypothetical protein NMA1602 [imported] - Neisseria meningitidis (strain Z2491 serogroup
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: GB1853
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holtroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: AB1775; MUID:20222556; PMID:10761919
A:Accession: GB1853
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-36 <PAR>
A:Cross-references: UNIPROT:Q9JTW3; GB:AL162756; GB:AL157959; NID:G7380091; PIDN:CAB8483
C:Genetics:
A:Experimental source: serogroup A, strain Z2491
A:Gene: NMA1602

Query Match 37.5%; Score 33; DB 2; Length 36;
Best Local Similarity 38.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 4 KAAELNKSKELE 16
: | : | : | : | :
Db 19 RTREINAKTKEIQ 31

RESULT 45
E882289
hypothetical protein VC0707 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

```

```

C.Accession: E82289
R.Heldberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gilm, M.L.; Dodson, R.J.;
charlson, D.; Emolaeva, M.D.; Vamathavan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Native 406, 477-483, 2000
A.Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A.Reference number: A82035; MUID:20406833; PMID:10952301
A.Accession: E82289
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-42 <HEI>
A.Cross-references: UNIPROT:Q9KU22; GB:AE004157; GB:AE003852; NID:g9655148; PIDN:AAF9387
A.Experimental source: serogroup O1, strain N16961; biotype El Tor
C.Gene: VC0707
A.Map position: 1

Query Match      37.5%; Score 33; DB 2; Length 42;
Best Local Similarity 50.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      7 ELNCKSKELEQQ 18
      ::|||::|||
Db      18 QINNSKQKEPQ 29

```

Search completed: August 29, 2005, 17:05:23
Job time : 40 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2005, 17:04:51 ; Search time 158 Seconds

(without alignments)
44.764 Million cell updates/sec

Title: US-10-031-289-1331
Perfect score: 88
Sequence: 1 PTKAARLNQSKRLBQ 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues

Total number of hits satisfying chosen parameters: 800413

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/prodata/2/pubppaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/prodata/2/pubppaa/PCRT_NEW_PUB.pep:*
- 3: /cgn2_6/prodata/2/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/prodata/2/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/prodata/2/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/prodata/2/pubppaa/PCRTUS_PUBCOMB.pep:*
- 7: /cgn2_6/prodata/2/pubppaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/prodata/2/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/prodata/2/pubppaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/prodata/2/pubppaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/prodata/2/pubppaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/prodata/2/pubppaa/US09C_NEW_PUB.pep:*
- 13: /cgn2_6/prodata/2/pubppaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/prodata/2/pubppaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/prodata/2/pubppaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/prodata/2/pubppaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/prodata/2/pubppaa/US10E_PUBCOMB.pep:*
- 18: /cgn2_6/prodata/2/pubppaa/US10F_NEW_PUB.pep:*
- 19: /cgn2_6/prodata/2/pubppaa/US11A_PUBCOMB.pep:*
- 20: /cgn2_6/prodata/2/pubppaa/US11_NEW_PUB.pep:*
- 21: /cgn2_6/prodata/2/pubppaa/US60_NEW_PUB.pep:*
- 22: /cgn2_6/prodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	42	47.7	35	10	US-09-259-658-2
2	42	47.7	35	10	US-09-259-658-26
3	42	47.7	35	10	US-09-259-658-28
4	42	47.7	35	10	US-09-259-658-31
5	42	47.7	35	10	US-09-259-658-46
6	42	47.7	35	10	US-09-259-658-47
7	42	47.7	35	10	US-09-259-658-51
8	42	47.7	35	10	US-09-259-658-53
9	42	47.7	35	10	US-09-259-658-55
10	42	47.7	35	10	US-09-259-658-57
11	42	47.7	43	10	US-09-259-658-27

12	42	47.7	43	10	US-09-259-658-40	Sequence 40, Appl
13	42	47.7	43	10	US-09-259-658-54	Sequence 54, Appl
14	42	47.7	72	15	US-10-424-593-201587	Sequence 201587,
15	41	46.6	83	15	US-10-424-599-180206	Sequence 180206,
16	41	46.6	89	14	US-10-059-720-64	Sequence 64, Appl
17	41	46.6	95	16	US-10-425-115-327771	Sequence 327771,
18	40	45.5	43	10	US-09-259-658-42	Sequence 42, Appl
19	40	45.5	49	16	US-10-425-115-192131	Sequence 192131,
20	40	45.5	65	15	US-10-424-599-257752	Sequence 257752,
21	40	45.5	69	16	US-10-425-115-346187	Sequence 346187,
22	40	45.5	77	15	US-10-424-599-231921	Sequence 231921,
23	40	45.5	83	16	US-10-425-115-292182	Sequence 292182,
24	39.5	44.9	46	10	US-09-898-554-29	Sequence 29, Appl
25	39.5	44.9	46	10	US-09-898-554-32	Sequence 32, Appl
26	39.5	44.9	46	10	US-09-898-554-34	Sequence 34, Appl
27	39.5	44.9	46	16	US-10-482-673-29	Sequence 29, Appl
28	39.5	44.9	46	16	US-10-482-673-32	Sequence 32, Appl
29	39.5	44.9	46	16	US-10-482-673-34	Sequence 34, Appl
30	39	44.3	25	14	US-10-071-174-27	Sequence 27, Appl
31	39	44.3	31	9	US-09-770-102A-37	Sequence 37, Appl
32	39	44.3	33	10	US-09-259-658-49	Sequence 49, Appl
33	39	44.3	34	14	US-10-059-720-53	Sequence 53, Appl
34	39	44.3	34	14	US-10-059-720-57	Sequence 57, Appl
35	39	44.3	35	14	US-10-161-205-42	Sequence 42, Appl
36	39	44.3	36	16	US-10-639-067-227	Sequence 227, App
37	39	44.3	36	16	US-10-639-067-228	Sequence 228, App
38	39	44.3	40	15	US-10-261-161-95	Sequence 95, Appl
39	39	44.3	43	10	US-09-259-658-1	Sequence 1, Appl
40	39	44.3	54	16	US-10-639-067-239	Sequence 239, App
41	39	44.3	57	16	US-10-425-115-251633	Sequence 251633,
42	39	44.3	79	16	US-10-437-963-161573	Sequence 161573,
43	39	44.3	92	16	US-10-425-115-243779	Sequence 243779,
44	39	44.3	92	16	US-10-425-115-328151	Sequence 328151,
45	38	43.2	26	14	US-10-059-720-3	Sequence 3, Appl
46	38	43.2	31	10	US-09-259-658-58	Sequence 58, Appl
47	38	43.2	33	14	US-10-059-720-38	Sequence 38, Appl
48	38	43.2	34	14	US-10-059-720-54	Sequence 54, Appl
49	38	43.2	35	10	US-09-259-658-39	Sequence 39, Appl
50	38	43.2	35	10	US-09-259-658-41	Sequence 41, Appl
51	38	43.2	35	10	US-09-259-658-43	Sequence 43, Appl
52	38	43.2	35	14	US-10-161-205-48	Sequence 48, Appl
53	38	43.2	63	15	US-10-424-599-217583	Sequence 217583,
54	38	43.2	73	15	US-10-276-774-1376	Sequence 1376, Ap
55	38	43.2	77	15	US-10-282-1220A-62270	Sequence 62270, A
56	38	43.2	79	14	US-10-177-725-16	Sequence 16, Appl
57	38	43.2	79	15	US-10-393-448-16	Sequence 16, Appl
58	38	43.2	83	14	US-10-177-725-12	Sequence 12, Appl
59	38	43.2	83	15	US-10-393-448-12	Sequence 12, Appl
60	38	43.2	84	11	US-09-864-408A-4954	Sequence 4954, Ap
61	38	43.2	84	14	US-10-059-720-11	Sequence 11, Appl
62	38	43.2	85	14	US-10-177-725-15	Sequence 15, Appl
63	38	43.2	85	15	US-10-393-448-15	Sequence 15, Appl
64	38	43.2	86	14	US-10-059-720-17	Sequence 17, Appl
65	38	43.2	86	14	US-10-059-720-52	Sequence 52, Appl
66	38	43.2	97	14	US-10-059-720-19	Sequence 19, Appl
67	37	42.0	46	10	US-09-898-554-31	Sequence 31, Appl
68	37	42.0	46	10	US-09-898-554-36	Sequence 36, Appl
69	37	42.0	46	10	US-09-898-554-37	Sequence 37, Appl
70	37	42.0	46	16	US-10-437-963-118713	Sequence 118713,
71	37	42.0	46	16	US-10-482-673-31	Sequence 31, Appl
72	37	42.0	46	16	US-10-482-673-36	Sequence 36, Appl
73	37	42.0	46	16	US-10-482-673-37	Sequence 37, Appl
74	37	42.0	55	15	US-10-424-599-189643	Sequence 189643,
75	37	42.0	56	16	US-10-437-963-120035	Sequence 120035,
76	37	42.0	57	9	US-09-864-761-41864	Sequence 41864, A
77	37	42.0	60	15	US-10-424-599-146989	Sequence 146989,
78	37	42.0	61	14	US-10-029-388-28948	Sequence 28948, A
79	37	42.0	66	16	US-10-767-701-34534	Sequence 34534, A
80	37	42.0	67	17	US-10-472-928-4582	Sequence 4582, Ap
81	37	42.0	68	16	US-10-425-115-318540	Sequence 318540,
82	37	42.0	72	15	US-10-335-977-5348	Sequence 5348, Ap
83	37	42.0	73	15	US-10-335-977-5347	Sequence 5347, Ap
84	37	42.0	74	18	US-10-426-124-11	Sequence 11, Appl

85	37	42.0	76	16	US-10-425-115-363599	Sequence 363599,
86	37	42.0	80	14	US-10-177-725-14	Sequence 14, Appl
87	37	42.0	80	15	US-10-393-449-14	Sequence 14, Appl
88	37	42.0	82	9	US-09-939-980-343	Sequence 343, App
89	37	42.0	90	15	US-10-424-599-231423	Sequence 231423,
90	37	42.0	90	16	US-10-425-115-333229	Sequence 333229,
91	36.5	41.5	82	16	US-10-425-115-232405	Sequence 232405,
92	36	40.9	20	17	US-10-690-276-709	Sequence 709, App
93	36	40.9	22	10	US-09-259-658-59	Sequence 59, Appl
94	36	40.9	23	14	US-10-059-720-2	Sequence 2, Appl1
95	36	40.9	33	14	US-10-181-654-6	Sequence 6, Appl1
96	36	40.9	33	14	US-10-181-654-26	Sequence 26, Appl
97	36	40.9	33	14	US-10-338-083-18	Sequence 18, Appl
98	36	40.9	33	16	US-10-611-399-18	Sequence 18, Appl
99	36	40.9	33	17	US-10-794-751-18	Sequence 18, Appl
100	36	40.9	33	17	US-10-944-473-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-259-658-2
; Sequence 2, Application US/09259658
; Publication No. US20030032054A1
; GENERAL INFORMATION:
; APPLICANT: Colyer
; APPLICANT: Craig
; APPLICANT: Maschio
; APPLICANT: Mezna
; TITLE OF INVENTION: Compositions And Methods For Monitoring The
; TITLE OF INVENTION: Modification State Of A Pair Of Polypeptides
; FILE REFERENCE: colyer 4256/79245
; CURRENT APPLICATION NUMBER: US/09/259, 658
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Protein binding
; OTHER INFORMATION: motif.
US-09-259-658-2

Query Match 47.7%; Score 42; DB 10; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 OKAAELNOKSKLEEQ 18
DB 6 OKIAQLKOKNAQLKOK 21
RESULT 2
US-09-259-658-26
; Sequence 26, Application US/09259658
; Publication No. US20030032054A1
; GENERAL INFORMATION:
; APPLICANT: Colyer
; APPLICANT: Craig
; APPLICANT: Maschio
; APPLICANT: Mezna
; TITLE OF INVENTION: Compositions And Methods For Monitoring The
; TITLE OF INVENTION: Modification State Of A Pair Of Polypeptides
; FILE REFERENCE: colyer 4256/79245
; CURRENT APPLICATION NUMBER: US/09/259, 658
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 26
; LENGTH: 35

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
; OTHER INFORMATION: peptide used to assay chymotrypsin activity.
US-09-259-658-26

Query Match 47.7%; Score 42; DB 10; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 OKAAELNOKSKLEEQ 18
DB 6 OKIAQLKOKNAQLKOK 21

RESULT 3
US-09-259-658-28
; Sequence 28, Application US/09259658
; Publication No. US20030032054A1
; GENERAL INFORMATION:
; APPLICANT: Colyer
; APPLICANT: Craig
; APPLICANT: Maschio
; APPLICANT: Mezna
; TITLE OF INVENTION: Compositions And Methods For Monitoring The
; TITLE OF INVENTION: Modification State Of A Pair Of Polypeptides
; FILE REFERENCE: colyer 4256/79245
; CURRENT APPLICATION NUMBER: US/09/259, 658
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 28
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
; OTHER INFORMATION: peptide used to assay thrombin activity.
US-09-259-658-28

Query Match 47.7%; Score 42; DB 10; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 OKAAELNOKSKLEEQ 18
DB 6 OKIAQLKOKNAQLKOK 21

RESULT 4
US-09-259-658-31
; Sequence 31, Application US/09259658
; Publication No. US20030032054A1
; GENERAL INFORMATION:
; APPLICANT: Colyer
; APPLICANT: Craig
; APPLICANT: Maschio
; APPLICANT: Mezna
; TITLE OF INVENTION: Compositions And Methods For Monitoring The
; TITLE OF INVENTION: Modification State Of A Pair Of Polypeptides
; FILE REFERENCE: colyer 4256/79245
; CURRENT APPLICATION NUMBER: US/09/259, 658
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 31
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
; OTHER INFORMATION: peptide used to assay TEV protease.

US-09-259-658-31

Query Match 47.7%; Score 42; DB 10; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 OKAAELNOKSKLEEQ 18
|||:|:|:|:|:|:|:
DB 6 OKIAQLKQKNAQLKQK 21

RESULT 5

US-09-259-658-46
; Sequence 46, Application US/09259658
; Publication No. US20030032054A1
; GENERAL INFORMATION:
; APPLICANT: Colyer
; APPLICANT: Craig
; APPLICANT: Maschio
; APPLICANT: Mezna

; TITLE OF INVENTION: Compositions And Methods For Monitoring The
; FILE REFERENCE: colyer 4256/79245
; CURRENT APPLICATION NUMBER: US/09/259,658
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 46
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide based upon p67/srt glycosylation acceptor
; OTHER INFORMATION: site.
US-09-259-658-46

Query Match 47.7%; Score 42; DB 10; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 OKAAELNOKSKLEEQ 18
|||:|:|:|:|:|:|:
DB 6 OKIAQLKQKNAQLKQK 21

RESULT 6

US-09-259-658-47
; Sequence 47, Application US/09259658
; Publication No. US20030032054A1
; GENERAL INFORMATION:
; APPLICANT: Colyer
; APPLICANT: Craig
; APPLICANT: Maschio
; APPLICANT: Mezna

; TITLE OF INVENTION: Compositions And Methods For Monitoring The
; FILE REFERENCE: colyer 4256/79245
; CURRENT APPLICATION NUMBER: US/09/259,658
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 47
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide used to assay chymotrypsin activity.
US-09-259-658-47

Query Match 47.7%; Score 42; DB 10; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;

Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 OKAAELNOKSKLEEQ 18
|||:|:~|:|:|:|:|:|:
DB 6 OKIAQLKQKNAQLKQK 21

RESULT 7

US-09-259-658-51
; Sequence 51, Application US/09259658
; Publication No. US20030032054A1
; GENERAL INFORMATION:
; APPLICANT: Colyer
; APPLICANT: Craig
; APPLICANT: Maschio
; APPLICANT: Mezna

; TITLE OF INVENTION: Compositions And Methods For Monitoring The
; FILE REFERENCE: colyer 4256/79245
; CURRENT APPLICATION NUMBER: US/09/259,658
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 51
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide used to assay exopeptidase activity.
US-09-259-658-51

Query Match 47.7%; Score 42; DB 10; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 OKAAELNOKSKLEEQ 18
|||:|:~|:|:|:|:|:|:
DB 6 OKIAQLKQKNAQLKQK 21

RESULT 8

US-09-259-658-53
; Sequence 53, Application US/09259658
; Publication No. US20030032054A1
; GENERAL INFORMATION:
; APPLICANT: Colyer
; APPLICANT: Craig
; APPLICANT: Maschio
; APPLICANT: Mezna

; TITLE OF INVENTION: Compositions And Methods For Monitoring The
; FILE REFERENCE: colyer 4256/79245
; CURRENT APPLICATION NUMBER: US/09/259,658
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 53
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide used to assay TEV protease.
US-09-259-658-53

Query Match 47.7%; Score 42; DB 10; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 OKAAELNOKSKLEEQ 18
|||:|:~|:|:~|:|:~|:~|:
DB 6 OKIAQLKQKNAQLKQK 21

```
RESULT 9
US-09-259-658-55
; Sequence 55, Application US/09259658
; Publication No. US20030032054A1
; GENERAL INFORMATION:
; APPLICANT: Colyer
; APPLICANT: Craig
; APPLICANT: Maschio
; APPLICANT: Mezna
; TITLE OF INVENTION: Compositions And Methods For Monitoring The
; TITLE OF INVENTION: Modification State Of A Pair Of Polypeptides
; FILE REFERENCE: colyer 4256/79245
; CURRENT APPLICATION NUMBER: US/09/259,658
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 55
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
; OTHER INFORMATION: peptide used to assay thrombin activity.
US-09-259-658-55
```

```
Query Match 47.7%; Score 42; DB 10; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 3 QKAAELNOKSKELRQ 18
Db 6 QKIAQLKQKNQQLKQK 21
```

```
RESULT 10
US-09-259-658-57
; Sequence 57, Application US/09259658
; Publication No. US20030032054A1
; GENERAL INFORMATION:
; APPLICANT: Colyer
; APPLICANT: Craig
; APPLICANT: Maschio
; APPLICANT: Mezna
; TITLE OF INVENTION: Compositions And Methods For Monitoring The
; TITLE OF INVENTION: Modification State Of A Pair Of Polypeptides
; FILE REFERENCE: colyer 4256/79245
; CURRENT APPLICATION NUMBER: US/09/259,658
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 57
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
; OTHER INFORMATION: peptide used to assay caspase activity.
US-09-259-658-57
```

```
Query Match 47.7%; Score 42; DB 10; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 3 QKAAELNOKSKELRQ 18
Db 6 QKIAQLKQKNQQLKQK 21
```

```
RESULT 11
US-09-259-658-27
; Sequence 27, Application US/09259658
```

```
; Publication No. US20030032054A1
; GENERAL INFORMATION:
; APPLICANT: Colyer
; APPLICANT: Craig
; APPLICANT: Maschio
; APPLICANT: Mezna
; TITLE OF INVENTION: Compositions And Methods For Monitoring The
; TITLE OF INVENTION: Modification State Of A Pair Of Polypeptides
; FILE REFERENCE: colyer 4256/79245
; CURRENT APPLICATION NUMBER: US/09/259,658
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 27
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
; OTHER INFORMATION: peptide used to assay thrombin activity.
US-09-259-658-27
```

```
Query Match 47.7%; Score 42; DB 10; Length 43;
Best Local Similarity 43.8%; Pred. No. 43;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 3 QKAAELNOKSKELRQ 18
Db 14 QETIAQLRQESALQLEQ 29
```

```
RESULT 12
US-09-259-658-40
; Sequence 40, Application US/09259658
; Publication No. US20030032054A1
; GENERAL INFORMATION:
; APPLICANT: Colyer
; APPLICANT: Craig
; APPLICANT: Maschio
; APPLICANT: Mezna
; TITLE OF INVENTION: Compositions And Methods For Monitoring The
; TITLE OF INVENTION: Modification State Of A Pair Of Polypeptides
; FILE REFERENCE: colyer 4256/79245
; CURRENT APPLICATION NUMBER: US/09/259,658
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 40
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
; OTHER INFORMATION: peptide used to assay CaMK-II activity.
US-09-259-658-40
```

```
Query Match 47.7%; Score 42; DB 10; Length 43;
Best Local Similarity 50.0%; Pred. No. 43;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 3 QKAAELNOKSKELRQ 18
Db 14 QETIAQLRQESALQLEQ 29
```

```
RESULT 13
US-09-259-658-54
; Sequence 54, Application US/09259658
; Publication No. US20030032054A1
; GENERAL INFORMATION:
; APPLICANT: Colyer
; APPLICANT: Craig
; APPLICANT: Maschio
```

APPLICANT: Mezna
TITLE OF INVENTION: Compositions And Methods For Monitoring The
FILE REFERENCE: colyer 4256/79245
CURRENT APPLICATION NUMBER: US/09/259,658
CURRENT FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 59
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 54
LENGTH: 43
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-259-658-54

Query Match 47.7%; Score 42; DB 10; Length 43;
Best Local Similarity 43.8%; Pred. No. 43;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Oy 3 QKAEINQSKLEEQ 18
Db 14 QEIAQLEQENRQLEQE 29

RESULT 14
US-10-424-599-201587
Sequence 201587, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 201587
LENGTH: 72
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_24059C.1.pdp
US-10-424-599-201587

Query Match 47.7%; Score 42; DB 15; Length 72;
Best Local Similarity 69.2%; Pred. No. 75;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 6 AELNQSKLEEQ 18
Db 15 AYLEKKELEEQ 27

RESULT 15
US-10-424-599-180206
Sequence 180206, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 180206
LENGTH: 83
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_133740C.1.pdp
US-10-424-599-180206

Query Match 46.6%; Score 41; DB 15; Length 83;
Best Local Similarity 56.2%; Pred. No. 1.2e+02;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 1 PTKAEINQSKLE 16
Db 25 PLAKPAFEDKKELE 40

RESULT 16
US-10-059-720-64
Sequence 64, Application US/10059720
Publication No. US20030027314A1
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN

NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS WORD 97
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/059,720
FILING DATE: 29-Jan-2002
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: 08/690,011
FILING DATE: 31-JULY-1996

ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849

INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 89 amino acids
TYPE: amino acid
STRANDEDNESS: unknown

MOLECULE TYPE: linear
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-10-059-720-64

Query Match 46.6%; Score 41; DB 14; Length 89;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Oy 3 QKAEINQSKLEEQ 18
::| | | | | | | | | |

Db 34 KEAELEQENAELEQ 49

RESULT 17

US-10-425-115-327771
; Sequence 327771, Application US/10425115
; Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 327771

LENGTH: 95

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: MRT4577_61996C.1.pep

US-10-425-115-327771

Query Match 46.6%; Score 41; DB 16; Length 95;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;

Matches 8; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 3 OKAAELNOKSELEQ 18

Db 41 RKAELEDRKAELEARK 56

RESULT 18

US-09-259-658-42

Sequence 42, Application US/09259658

Publication No. US20030032054A1

GENERAL INFORMATION:

APPLICANT: Colyer

APPLICANT: Craig

APPLICANT: Maschio

APPLICANT: Mezna

TITLE OF INVENTION: Compositions And Methods For Monitoring The

FILE REFERENCE: colyer 4256/79245

CURRENT APPLICATION NUMBER: US/09/259,658

CURRENT FILING DATE: 1999-02-26

NUMBER OF SEQ ID NOS: 59

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 42

LENGTH: 43

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: peptide used to assay S6 kinase activity.

US-09-259-658-42

Query Match 45.5%; Score 40; DB 10; Length 43;
Best Local Similarity 50.0%; Pred. No. 86;

Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 OKAAELNOKSELEQ 18

Db 14 QEIARLRQESAQLQDE 29

RESULT 19

US-10-425-115-192131

Sequence 192131, Application US/10425115

Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 192131

LENGTH: 49

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: MRT4577_106808C.1.pep

US-10-425-115-192131

Query Match 45.5%; Score 40; DB 16; Length 49;
Best Local Similarity 60.0%; Pred. No. 99;

Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 OKAAELNOKSELEQ 17

Db 25 QMADQVAKSSELEQ 39

RESULT 20

US-10-424-599-257752

Sequence 257752, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J

APPLICANT: Kovalic, David K

APPLICANT: Zhou, Yihua

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 257752

LENGTH: 65

TYPE: PRT

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_74774C.1.pep

US-10-424-599-257752

Query Match 45.5%; Score 40; DB 15; Length 65;
Best Local Similarity 53.3%; Pred. No. 1.3e+02;

Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 PTOAAELNOKSELE 15

Db 29 PLOAAELVQKTORI 43

RESULT 21

US-10-425-115-346187

Sequence 346187, Application US/10425115

Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 346187
LENGTH: 69
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_78889C.1.pep
US-10-425-115-346187

Query Match 45.5%; Score 40; DB 16; Length 69;
Best Local Similarity 43.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 3 OKAAELNOKSKELEQ 18
Db 24 KKOELNRKAEELDR 39

RESULT 22
US-10-424-599-231921
Sequence 231921, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Zhou Yihua
APPLICANT: Kovalic David K
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT FILING DATE: 2003-04-28
CURRENT APPLICATION NUMBER: US/10/424,599
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 231921
LENGTH: 77
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_51447C.1.pep
US-10-424-599-231921

Query Match 45.5%; Score 40; DB 15; Length 77;
Best Local Similarity 47.1%; Pred. No. 1.6e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 2 TOKAAELNOKSKELEQ 18
Db 17 TASTDELKDKKEIEBQ 33

RESULT 23
US-10-425-115-292182
Sequence 292182, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 292182
LENGTH: 83
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_29556C.1.pep
US-10-425-115-292182

Query Match 45.5%; Score 40; DB 16; Length 83;
Best Local Similarity 38.9%; Pred. No. 1.8e+02;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQ 18
Db 7 PNOGNIETNRKSPQMEQK 24

RESULT 24
US-09-898-554-29
Sequence 29, Application US/09898554
Publication No. US20030068673A1
GENERAL INFORMATION:
APPLICANT: TALL, ALAN R
APPLICANT: WELCH, CARIE L
APPLICANT: LIANG, CHIEN-PING
TITLE OF INVENTION: ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 1 (ATHSQ1) AND ATHEROS
TITLE OF INVENTION: SUSCEPTIBILITY GENE LOCUS 2 (ATHSQ2)
FILE REFERENCE: 0575/64077
CURRENT APPLICATION NUMBER: US/09/898,554
CURRENT FILING DATE: 2001-07-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn version 3.1
SEQ ID NO 29
LENGTH: 46
TYPE: PRT
ORGANISM: Murinae gen. sp.
FEATURE:
NAME/KEY: MISC FEATURE
OTHER INFORMATION: ISOFORM 1 REPEAT #1
US-09-898-554-29

Query Match 44.9%; Score 39.5; DB 10; Length 46;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

Qy 2 TOKAAELNOKSKELEQ 17
Db 14 TOK---LNEKSKOE 26

RESULT 25
US-09-898-554-32
Sequence 32, Application US/09898554
Publication No. US20030068673A1
GENERAL INFORMATION:
APPLICANT: TALL, ALAN R
APPLICANT: WELCH, CARIE L
APPLICANT: LIANG, CHIEN-PING
TITLE OF INVENTION: ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 1 (ATHSQ1) AND ATHEROS
TITLE OF INVENTION: SUSCEPTIBILITY GENE LOCUS 2 (ATHSQ2)
FILE REFERENCE: 0575/64077
CURRENT APPLICATION NUMBER: US/09/898,554
CURRENT FILING DATE: 2001-07-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn version 3.1
SEQ ID NO 32
LENGTH: 46
TYPE: PRT
ORGANISM: Murinae gen. sp.
FEATURE:
NAME/KEY: MISC FEATURE
OTHER INFORMATION: ISOFORM 3 REPEAT #1
US-09-898-554-32

Query Match 44.9%; Score 39.5; DB 10; Length 46;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

Qy 2 TOKAAELNOKSKELEQ 17
Db 14 TOK---LNEKSKOE 26

RESULT 25
US-09-898-554-32
Sequence 32, Application US/09898554
Publication No. US20030068673A1
GENERAL INFORMATION:
APPLICANT: TALL, ALAN R
APPLICANT: WELCH, CARIE L
APPLICANT: LIANG, CHIEN-PING
TITLE OF INVENTION: ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 1 (ATHSQ1) AND ATHEROS
TITLE OF INVENTION: SUSCEPTIBILITY GENE LOCUS 2 (ATHSQ2)
FILE REFERENCE: 0575/64077
CURRENT APPLICATION NUMBER: US/09/898,554
CURRENT FILING DATE: 2001-07-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn version 3.1
SEQ ID NO 32
LENGTH: 46
TYPE: PRT
ORGANISM: Murinae gen. sp.
FEATURE:
NAME/KEY: MISC FEATURE
OTHER INFORMATION: ISOFORM 3 REPEAT #1
US-09-898-554-32

Query Match 44.9%; Score 39.5; DB 10; Length 46;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

```
Db      14 TOK---LNEKSKEOE 26

RESULT 26
US-09-898-554-34
; Sequence 34, Application US/09898554
; Publication No. US20030068673A1
; GENERAL INFORMATION:
; APPLICANT: TALL, ALAN R
; APPLICANT: WELCH, CARIE L
; APPLICANT: LIANG, CHEN-PING
; TITLE OF INVENTION: ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 1 (ATHSQ1) AND ATHEROS
; TITLE OF INVENTION: SUSCEPTIBILITY GENE LOCUS 2 (ATHSQ2)
; FILE REFERENCE: 0575/64077
; CURRENT APPLICATION NUMBER: US/09/898,554
; CURRENT FILING DATE: 2001-07-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: ISOFORM 4 REPEAT #1
US-09-898-554-34

Query Match      44.9%; Score 39.5; DB 16; Length 46;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY      2 TOKAELNOKSKELEQ 17
      ||| ||:|||||:
      14 TOK---LNEKSKEOE 26

RESULT 27
US-10-482-673-29
; Sequence 29, Application US/10482673
; Publication No. US20040267000A1
; GENERAL INFORMATION:
; APPLICANT: THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK
; TITLE OF INVENTION: ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 1 (ATHSQ1) AND ATHEROSC
; TITLE OF INVENTION: SUSCEPTIBILITY GENE LOCUS 2 (ATHSQ2)
; FILE REFERENCE: 0575/64077apct
; CURRENT APPLICATION NUMBER: US/10/482,673
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 09/898,554
; PRIOR FILING DATE: 2001-07-02
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: ISOFORM 1 REPEAT #1
US-10-482-673-29

Query Match      44.9%; Score 39.5; DB 16; Length 46;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY      2 TOKAELNOKSKELEQ 17
      ||| ||:|||||:
      14 TOK---LNEKSKEOE 26

RESULT 28
US-10-482-673-32
; Sequence 32, Application US/10482673
; Publication No. US20040267000A1
; GENERAL INFORMATION:
; APPLICANT: THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK
; TITLE OF INVENTION: ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 1 (ATHSQ1) AND ATHEROSC
; TITLE OF INVENTION: SUSCEPTIBILITY GENE LOCUS 2 (ATHSQ2)
; FILE REFERENCE: 0575/64077apct
; CURRENT APPLICATION NUMBER: US/10/482,673
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 09/898,554
; PRIOR FILING DATE: 2001-07-02
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: ISOFORM 3 REPEAT #1
US-10-482-673-32

Query Match      44.9%; Score 39.5; DB 16; Length 46;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY      2 TOKAELNOKSKELEQ 17
      ||| ||:|||||:
      14 TOK---LNEKSKEOE 26

RESULT 29
US-10-482-673-34
; Sequence 34, Application US/10482673
; Publication No. US20040267000A1
; GENERAL INFORMATION:
; APPLICANT: THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK
; TITLE OF INVENTION: ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 1 (ATHSQ1) AND ATHEROSC
; TITLE OF INVENTION: SUSCEPTIBILITY GENE LOCUS 2 (ATHSQ2)
; FILE REFERENCE: 0575/64077apct
; CURRENT APPLICATION NUMBER: US/10/482,673
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 09/898,554
; PRIOR FILING DATE: 2001-07-02
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: ISOFORM 4 REPEAT #1
US-10-482-673-34

Query Match      44.9%; Score 39.5; DB 16; Length 46;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY      2 TOKAELNOKSKELEQ 17
      ||| ||:|||||:
      14 TOK---LNEKSKEOE 26

RESULT 30
US-10-071-174-27
; Sequence 27, Application US/10071174
; Publication No. US20030176671A1
; GENERAL INFORMATION:
; APPLICANT: REED, JOHN C.
; APPLICANT: KE, NING
; APPLICANT: GODZIK, ADAM
; TITLE OF INVENTION: APOPTOSIS MODULATOR BCL-B AND METHODS FOR MAKING AND
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 087102-0272558
```

3 OKAAEINOKSKELEO 17

Db 20 KEAELEQENAELEQ 34

RESULT 34
US-10-059-720-57
; Sequence 57, Application US/10059720
; Publication No. US20030027314A1
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; KRYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS WORD 97
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/059,720
; FILING DATE: 29-Jan-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; APPLICATION NUMBER: 08/690,011
; FILING DATE: 31-JULY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-419US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)751-6849
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 57:
US-10-059-720-57

Query Match 44.3%; Score 39; DB 14; Length 34;
Best Local Similarity 53.3%; Pred. No. 94;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 OKAAELNQSKKELEQ 17
::|||::|||
Db 20 KEAELEQENAELEQ 34

RESULT 35
US-10-161-205-42
; Sequence 42, Application US/10161205
; Publication No. US20030100037A1
; GENERAL INFORMATION:
; APPLICANT: Cyclacel
; APPLICANT: Colyer, John
; TITLE OF INVENTION: Chemical Modification
; FILE REFERENCE: 10069/1094
; CURRENT APPLICATION NUMBER: US/10/161,205
; CURRENT FILING DATE: 2002-06-03

; PRIOR APPLICATION NUMBER: US 09/470,485
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: US 09/146,549
; PRIOR FILING DATE: 1998-09-03
; PRIOR APPLICATION NUMBER: PCT/GB98/02565
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: GB 9718358.6
; PRIOR FILING DATE: 1997-08-30
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthesized peptide
US-10-161-205-42

Query Match 44.3%; Score 39; DB 14; Length 35;
Best Local Similarity 43.8%; Pred. No. 97;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 3 OKAAELNQSKKELEQ 18
|||::|||
Db 13 QEIAPLEQENAELEQ 28

RESULT 36
US-10-639-067-227
; Sequence 227, Application US/10639067
; Publication No. US20040214255A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Incorporated
; APPLICANT: Heichman, Karen
; APPLICANT: Bartel, Paul
; APPLICANT: Sugiyama, Janice
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING DIABETES
; FILE REFERENCE: 1833.01
; CURRENT APPLICATION NUMBER: US/10/639,067
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US 09/556,941
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: US 60/130,389
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: US 60/140,693
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: US 60/156,947
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US 60/163,073
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/168,378
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: US 60/168,376
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 227
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-639-067-227

Query Match 44.3%; Score 39; DB 16; Length 36;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 OKAAELNQSKKELEQ 18
|||::|||
Db 6 QKTELKQKLAIDLQKQ 21

RESULT 37
US-10-639-067-228

```
Sequence 228, Application US/10639067
; Publication No. US20040214255A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Incorporated
; APPLICANT: Heichman, Karen
; APPLICANT: Bartel, Paul
; APPLICANT: Sugiyama, Janice
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING DIABETES
; FILE REFERENCE: 1833.01
; CURRENT APPLICATION NUMBER: US/10/639,067
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US 09/556,941
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: US 60/130,389
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: US 60/140,693
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: US 60/156,947
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US 60/163,073
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/168,378
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: US 60/168,376
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 228
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-639-067-228
```

```
Query Match          44.3%; Score 39; DB 16; Length 36;
Best Local Similarity 52.9%; Pred. No. 1e+02;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy      2  TOKAAELNOKSKELFQQ 18
Db      19 TAAEKEIKOKVKYLODQ 35
```

```
RESULT 38
US-10-261-161-95
; Sequence 95, Application US/10261161
; Publication No. US20040072270A1
; GENERAL INFORMATION:
; APPLICANT: Fernandez-Salas, Ester
; APPLICANT: Steward, Lance E.
; APPLICANT: Aoki, Kei Roger
; TITLE OF INVENTION: Cell-Based Fluorescence Resonance Energy
; TITLE OF INVENTION: Transfer (FRET) Assays For Clostridial Toxins
; FILE REFERENCE: P-AR 4804
; CURRENT APPLICATION NUMBER: US/10/261,161
; CURRENT FILING DATE: 2002-09-27
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 95
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-261-161-95
```

```
Query Match          44.3%; Score 39; DB 15; Length 40;
Best Local Similarity 46.7%; Pred. No. 1.1e+02;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy      3  OKAAELNOKSKELFQ 17
Db      8  QKSELGRADQLEQ 22
```

RESULT 39

```
US-09-259-658-1
; Sequence 1, Application US/09259658
; Publication No. US20030032054A1
; GENERAL INFORMATION:
; APPLICANT: Colyer
; APPLICANT: Craig
; APPLICANT: Maschio
; APPLICANT: Mezna
; TITLE OF INVENTION: Compositions And Methods For Monitoring The
; TITLE OF INVENTION: Modification State Of A Pair Of Polypeptides
; FILE REFERENCE: Colyer 4256/79245
; CURRENT APPLICATION NUMBER: US/09/259,658
; CURRENT FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Protein binding
; OTHER INFORMATION: motif.
US-09-259-658-1
```

```
Query Match          44.3%; Score 39; DB 10; Length 43;
Best Local Similarity 43.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy      3  OKAAELNOKSKELFQQ 18
Db      14 QETIAQLQENVAQLQEQE 29
```

```
RESULT 40
US-10-639-067-239
; Sequence 239, Application US/10639067
; Publication No. US20040214255A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Incorporated
; APPLICANT: Heichman, Karen
; APPLICANT: Bartel, Paul
; APPLICANT: Sugiyama, Janice
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING DIABETES
; FILE REFERENCE: 1833.01
; CURRENT APPLICATION NUMBER: US/10/639,067
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: US 09/556,941
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: US 60/130,389
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: US 60/140,693
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: US 60/156,947
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US 60/163,073
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/168,378
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: US 60/168,376
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 239
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-639-067-239
```

```
Query Match          44.3%; Score 39; DB 16; Length 54;
Best Local Similarity 52.9%; Pred. No. 1.6e+02;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy      2  TOKAAELNOKSKELFQQ 18
```

Db 23 TAAKELRQKVKYLQDQ 39

RESULT 41

US-10-425-115-251633
; Sequence 251633, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 251633
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(57)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_161075C.1.pep
US-10-425-115-251633

Query Match 44.3%; Score 39; DB 16; Length 57;
Best Local Similarity 61.5%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 KAAELNOKSKEL 16
:|:|:|:|:|:|
Db 2 EAGFNAXKKELE 14

RESULT 42

US-10-437-963-161573
; Sequence 161573, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 161573
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(79)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_60746C.1.pep
US-10-437-963-161573

Query Match 44.3%; Score 39; DB 16; Length 79;
Best Local Similarity 63.6%; Pred. No. 2.4e+02;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 PTOKAELNOK 11
||:|:|:|:|:|
Db 55 PTRKEVEVNOK 65

RESULT 43

US-10-425-115-243779
; Sequence 243779, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 243779
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_15390C.1.pep
US-10-425-115-243779

Query Match 44.3%; Score 39; DB 16; Length 92;
Best Local Similarity 61.5%; Pred. No. 2.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 OKAAELNOKSKEL 15
:|:|:|:|:|:|
Db 66 EKIAEANQKLEK 78

RESULT 44

US-10-425-115-328151
; Sequence 328151, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 328151
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_62338C.1.pep
US-10-425-115-328151

Query Match 44.3%; Score 39; DB 16; Length 92;
Best Local Similarity 53.3%; Pred. No. 2.8e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 OKAAELNOKSKELQ 17
:|:|:|:|:|:|
Db 78 KKAVMIQKSKELNQ 92

RESULT 45
US-10-059-720-3

```

: Sequence 3, Application US/10059720
: Publication No. US20030027314A1
: GENERAL INFORMATION:
: APPLICANT: VINSON, Charles R.
: KRYLOV, Dmitry
: TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
: CELLULAR PROTEIN
: NUMBER OF SEQUENCES: 64
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
: STREET: 345 Park Avenue
: CITY: New York
: STATE: NY
: COUNTRY: USA
: ZIP: 10154-0053
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: MS WORD 97
: SOFTWARE: ASCII
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/059,720
: FILING DATE: 29-Jan-2002
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/001,654
: FILING DATE: 31-JUL-1995
: APPLICATION NUMBER: 60/018,496
: FILING DATE: 29-MAY-1996
: APPLICATION NUMBER: 08/690,011
: FILING DATE: 31-JULY-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Serunian, Leslie A.
: REGISTRATION NUMBER: 35,353
: REFERENCE/DOCKET NUMBER: 2026-4199US2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212)758-4800
: TELEFAX: (212)751-6849
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 26 amino acids
: TYPE: amino acid
: STRANDEDNESS: unknown
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-059-720-3

```

```

Query Match          43.2% Score 38; DB 14; Length 26;
Best Local Similarity 43.8% Pred. No. 99;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
QY      3 QKAALNQSKLECO 18
       ::|||::|||::
Db      6 KEAELEQENALEEE 21

```

Search completed: August 29, 2005, 17:18:23
 Job time : 159 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2005, 16:59:05 ; Search time 42 seconds
(without alignments)
31.992 Million cell updates/sec

Title: US-10-031-289-1331
Perfect score: 88
Sequence: 1 PTKAAELINKSKLEEQ 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 324380

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.dep: *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.dep: *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.dep: *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.dep: *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.dep: *
6: /cgn2_6/ptodata/1/iaa/backfilest.dep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	50.0	43	3	US-09-015-030-6
2	41	46.6	89	2	US-08-690-011A-52
3	41	46.6	89	3	US-09-299-495F-64
4	40	45.5	82	4	US-09-513-999C-5870
5	39.5	44.9	46	4	US-09-898-554-29
6	39.5	44.9	46	4	US-09-898-554-32
7	39.5	44.9	46	4	US-09-898-554-34
8	39.5	44.9	27	4	US-09-544-664B-26
9	39	44.3	34	2	US-08-690-011A-53
10	39	44.3	34	2	US-08-690-011A-57
11	39	44.3	34	2	US-09-299-495F-53
12	39	44.3	34	3	US-09-299-495F-57
13	39	44.3	82	4	US-09-621-976-5189
14	39	44.3	98	2	US-08-690-011A-19
15	38	43.2	26	2	US-08-690-011A-3
16	38	43.2	26	3	US-09-299-495F-3
17	38	43.2	33	2	US-08-690-011A-38
18	38	43.2	33	3	US-09-299-495F-38
19	38	43.2	34	2	US-08-690-011A-54
20	38	43.2	34	3	US-09-299-495F-54
21	38	43.2	84	3	US-09-299-495F-11
22	38	43.2	85	4	US-09-134-000C-4168
23	38	43.2	86	3	US-09-299-495F-17
24	38	43.2	86	3	US-09-299-495F-52
25	38	43.2	88	2	US-08-690-011A-11
26	38	43.2	97	3	US-09-299-495F-19
27	37	42.0	42	3	US-08-956-307B-2

28	37	42.0	43	3	US-09-015-030-7	Sequence 7, Appl
29	37	42.0	46	4	US-09-898-554-31	Sequence 31, Appl
30	37	42.0	46	4	US-09-898-554-36	Sequence 36, Appl
31	37	42.0	46	4	US-09-898-554-37	Sequence 37, Appl
32	37	42.0	52	4	US-09-513-999C-5756	Sequence 5756, Ap
33	37	42.0	67	4	US-09-583-110-4322	Sequence 4322, Ap
34	37	42.0	74	4	US-09-490-291-11	Sequence 11, Appl
35	37	42.0	82	3	US-08-936-165A-343	Sequence 343, App
36	37	42.0	84	1	US-08-452-592B-7	Sequence 7, Appl
37	37	42.0	84	1	US-08-452-592B-8	Sequence 8, Appl
38	37	42.0	84	1	US-08-452-592B-9	Sequence 9, Appl
39	37	42.0	87	4	US-09-513-999C-5991	Sequence 5991, Ap
40	37	42.0	98	4	US-09-107-532A-5271	Sequence 5271, Ap
41	36.5	41.5	97	4	US-09-902-540-12287	Sequence 12287, A
42	36	40.9	25	2	US-08-690-011A-2	Sequence 2, Appl
43	36	40.9	25	3	US-09-299-495F-2	Sequence 2, Appl
44	36	40.9	79	4	US-09-640-211A-622	Sequence 622, App
45	36	40.9	83	4	US-09-640-211A-2181	Sequence 2181, Ap
46	36	40.9	96	4	US-09-328-352-5191	Sequence 5191, Ap
47	35.5	40.3	55	4	US-09-270-767-57475	Sequence 57475, A
48	35.5	40.3	79	4	US-09-248-796A-23855	Sequence 23855, A
49	35	39.8	46	4	US-09-898-554-38	Sequence 38, Appl
50	35	39.8	74	4	US-09-270-767-32762	Sequence 32762, A
51	35	39.8	74	4	US-09-270-767-47979	Sequence 47979, A
52	35	39.8	75	4	US-09-513-999C-7049	Sequence 7049, Ap
53	35	39.8	88	4	US-09-621-976-5903	Sequence 5903, Ap
54	35	39.8	88	4	US-09-621-976-5913	Sequence 5913, Ap
55	34	38.6	16	1	US-08-126-564A-36	Sequence 36, Appl
56	34	38.6	16	1	PCT-US94-09143-36	Sequence 36, Appl
57	34	38.6	19	2	US-08-690-011A-36	Sequence 36, Appl
58	34	38.6	19	3	US-09-299-495F-36	Sequence 36, Appl
59	34	38.6	22	6	5395759-4	Patent No. 5395759
60	34	38.6	22	6	5395759-4	Patent No. 5395759
61	34	38.6	24	2	US-08-690-011A-1	Sequence 1, Appl
62	34	38.6	24	3	US-09-299-495F-1	Sequence 1, Appl
63	34	38.6	26	2	US-08-690-011A-37	Sequence 37, Appl
64	34	38.6	26	3	US-09-299-495F-37	Sequence 37, Appl
65	34	38.6	34	1	US-07-956-700B-7	Sequence 7, Appl
66	34	38.6	34	1	US-08-475-537-7	Sequence 7, Appl
67	34	38.6	34	1	US-08-485-607-7	Sequence 7, Appl
68	34	38.6	34	2	US-08-475-879-7	Sequence 7, Appl
69	34	38.6	34	3	US-09-433-043B-7	Sequence 7, Appl
70	34	38.6	38	1	US-08-346-293-24	Sequence 24, Appl
71	34	38.6	42	1	US-08-346-293-25	Sequence 25, Appl
72	34	38.6	57	4	US-09-270-767-58256	Sequence 58256, A
73	34	38.6	64	4	US-09-205-258-580	Sequence 580, App
74	34	38.6	79	4	US-09-328-352-7073	Sequence 7073, Ap
75	34	38.6	79	4	US-09-732-210-1188	Sequence 1188, Ap
76	34	38.6	79	4	US-09-248-796A-25611	Sequence 25611, A
77	34	38.6	85	4	US-09-248-796A-27307	Sequence 27307, A
78	34	38.6	88	2	US-08-690-011A-27	Sequence 27, Appl
79	34	38.6	90	4	US-09-513-999C-5383	Sequence 5383, Ap
80	34	38.6	100	4	US-09-116-492A-28	Sequence 28, Appl
81	33	37.5	24	2	US-08-455-968B-23	Sequence 23, Appl
82	33	37.5	26	4	US-09-403-861A-7	Sequence 7, Appl
83	33	37.5	29	4	US-08-894-139-13	Sequence 13, Appl
84	33	37.5	29	4	US-08-894-139-14	Sequence 14, Appl
85	33	37.5	32	4	US-09-491-614B-31	Sequence 31, Appl
86	33	37.5	35	3	US-08-774-619B-13	Sequence 13, Appl
87	33	37.5	36	4	US-09-270-767-57518	Sequence 57518, A
88	33	37.5	42	3	US-08-956-307B-1	Sequence 1, Appl
89	33	37.5	59	4	US-09-513-999C-6526	Sequence 6526, Ap
90	33	37.5	59	4	US-09-513-999C-6778	Sequence 6778, Ap
91	33	37.5	60	4	US-09-270-767-60389	Sequence 60389, A
92	33	37.5	63	4	US-08-983-157B-4	Sequence 4, Appl
93	33	37.5	66	4	US-09-732-210-824	Sequence 824, App
94	33	37.5	68	4	US-09-621-976-7373	Sequence 7373, Ap
95	33	37.5	74	3	US-09-286-959B-19	Sequence 19, Appl
96	33	37.5	76	3	US-08-956-307B-17	Sequence 17, Appl
97	33	37.5	76	3	US-08-956-307B-18	Sequence 18, Appl
98	33	37.5	80	2	US-08-710-749-18	Sequence 18, Appl
99	33	37.5	80	4	US-09-147-875A-19	Sequence 19, Appl
100	33	37.5	81	4	US-09-621-976-6626	Sequence 6626, Ap

ALIGNMENTS

RESULT 1

US-09-015-030-6
Sequence 6, Application US/09015030
Patent No. 630317
GENERAL INFORMATION:
APPLICANT: Alber, Thomas C.
APPLICANT: Sharma, Victoria A.
APPLICANT: Nautiyal, Shiviati
TITLE OF INVENTION: Peptide Probes and Methods for Making the Same
FILE REFERENCE: A-64988-1/RFT/TNL
CURRENT APPLICATION NUMBER: US/09/015,030
CURRENT FILING DATE: 1998-01-28
PRIOR APPLICATION NUMBER: 60/036,219
PRIOR FILING DATE: 1997-01-28
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 43
TYPE: PRT
ORGANISM: Yeast
US-09-015-030-6

Query Match 50.0%; Score 44; DB 3; Length 43;
Best Local Similarity 50.0%; Pred. No. 4.5;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 3 OKAAELNOKSKELDEQ 18
Db 14 EYELKKNKDLDE 29

RESULT 2

US-08-690-011A-52
Sequence 52, Application US/08690011A
Patent No. 594243
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
TITLE OF INVENTION: CELLULAR PROTEIN
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,011A
FILING DATE: 31-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US1
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 89 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-690-011A-52

Query Match 46.6%; Score 41; DB 2; Length 89;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 3 OKAAELNOKSKELDEQ 18
Db 34 KEAELEQENAELEDE 49

RESULT 3

US-09-299-495F-64
Sequence 64, Application US/09299495F
Patent No. 6361968
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS WORD 97
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,495F
FILING DATE: 26-Apr-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: 08/690,011
FILING DATE: 31-JULY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 89 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-09-299-495F-64

Query Match 46.6%; Score 41; DB 3; Length 89;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Oy 3 OKAAELNOKSELEQ 18
 Db 34 KEAELEQENAELEQ 49

RESULT 4

US-09-513-999C-5870
 ; Sequence 5870, Application US/09513999C
 ; Patent No. 6783961
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, J.B.
 ; APPLICANT: Duclet, A.
 ; APPLICANT: Giordano, J.Y.
 ; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
 ; Patent No. 6783961
 ; FILE REFERENCE: 59.US2.REG
 ; CURRENT APPLICATION NUMBER: US/09/513.999C
 ; CURRENT FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/122,487
 ; PRIOR FILING DATE: 1999-02-26
 ; NUMBER OF SEQ ID NOS: 36681
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 5870
 ; LENGTH: 82
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-513-999C-5870

Query Match 45.5%; Score 40; DB 4; Length 82;
 Best Local Similarity 43.8%; Pred. No. 37;
 Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Oy 3 OKAAELNOKSELEQ 18
 Db 60 KKOELNKAELDR 75

RESULT 5

US-09-898-554-29
 ; Sequence 29, Application US/09898554
 ; Patent No. 6756228
 ; GENERAL INFORMATION:
 ; APPLICANT: TALL, ALAN R
 ; APPLICANT: WELCH, CARIE L
 ; APPLICANT: LIANG, CHEN-PING
 ; TITLE OF INVENTION: ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 1 (ATHSQ1) AND ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 2 (ATHSQ2)
 ; FILE REFERENCE: 0575/64077
 ; CURRENT APPLICATION NUMBER: US/09/898.554
 ; CURRENT FILING DATE: 2001-07-02
 ; NUMBER OF SEQ ID NOS: 40
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 29
 ; LENGTH: 46
 ; TYPE: PRT
 ; ORGANISM: Murinae gen. sp.
 ; FEATURE:
 ; NAME/KEY: MISC_FEATURE
 ; OTHER INFORMATION: ISOFORM 1 REPEAT #1
 US-09-898-554-29

Query Match 44.9%; Score 39.5; DB 4; Length 46;
 Best Local Similarity 62.5%; Pred. No. 24;
 Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

Oy 2 TOKAAELNOKSELEQ 17
 Db 14 TOK--LNEKSKOE 26

RESULT 6
 US-09-898-554-32

; Sequence 32, Application US/09898554
 ; Patent No. 6756228
 ; GENERAL INFORMATION:
 ; APPLICANT: TALL, ALAN R
 ; APPLICANT: WELCH, CARIE L
 ; APPLICANT: LIANG, CHEN-PING
 ; TITLE OF INVENTION: ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 1 (ATHSQ1) AND ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 2 (ATHSQ2)
 ; FILE REFERENCE: 0575/64077
 ; CURRENT APPLICATION NUMBER: US/09/898.554
 ; CURRENT FILING DATE: 2001-07-02
 ; NUMBER OF SEQ ID NOS: 40
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 32
 ; LENGTH: 46
 ; TYPE: PRT
 ; ORGANISM: Murinae gen. sp.
 ; FEATURE:
 ; NAME/KEY: MISC_FEATURE
 ; OTHER INFORMATION: ISOFORM 3 REPEAT #1
 US-09-898-554-32

Query Match 44.9%; Score 39.5; DB 4; Length 46;
 Best Local Similarity 62.5%; Pred. No. 24;
 Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

Oy 2 TOKAAELNOKSELEQ 17
 Db 14 TOK--LNEKSKOE 26

RESULT 7
 US-09-898-554-34

; Sequence 34, Application US/09898554
 ; Patent No. 6756228
 ; GENERAL INFORMATION:
 ; APPLICANT: TALL, ALAN R
 ; APPLICANT: WELCH, CARIE L
 ; APPLICANT: LIANG, CHEN-PING
 ; TITLE OF INVENTION: ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 1 (ATHSQ1) AND ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 2 (ATHSQ2)
 ; FILE REFERENCE: 0575/64077
 ; CURRENT APPLICATION NUMBER: US/09/898.554
 ; CURRENT FILING DATE: 2001-07-02
 ; NUMBER OF SEQ ID NOS: 40
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 34
 ; LENGTH: 46
 ; TYPE: PRT
 ; ORGANISM: Murinae gen. sp.
 ; FEATURE:
 ; NAME/KEY: MISC_FEATURE
 ; OTHER INFORMATION: ISOFORM 4 REPEAT #1
 US-09-898-554-34

Query Match 44.9%; Score 39.5; DB 4; Length 46;
 Best Local Similarity 62.5%; Pred. No. 24;
 Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

Oy 2 TOKAAELNOKSELEQ 17
 Db 14 TOK--LNEKSKOE 26

RESULT 8
 US-09-544-664B-26

; Sequence 26, Application US/09544664B
 ; Patent No. 6713280
 ; GENERAL INFORMATION:
 ; APPLICANT: Huang, Ziwei
 ; APPLICANT: Wang, Jialun
 ; APPLICANT: Zhang, Zhijia
 ; APPLICANT: Shan, Simei

APPLICANT: Lu, Zhixian
TITLE OF INVENTION: Enhancement of Peptide Cellular Uptake
FILE REFERENCE: 8321-68
CURRENT APPLICATION NUMBER: US/09/544,664B
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09352
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/128,202
PRIOR FILING DATE: 1999-04-07
NUMBER OF SEQ ID NOS: 58
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO: 26
LENGTH: 27
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence: Peptide
OTHER INFORMATION: segment from Bcl-2 superfamily
OTHER INFORMATION: polypeptide
US-09-544-664B-26

Query Match 44.3%; Score 39; DB 4; Length 27;
Best Local Similarity 44.4%; Pred. No. 16;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 PTOKAELNOKSKELEQ 18
|:||||:|:||||:
Db 5 PSATPAELRRAAELERR 22

RESULT 9

US-08-690-011A-53
Sequence 53, Application US/08690011A
Patent No. 5942433
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
TITLE OF INVENTION: CELLULAR PROTEIN
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,011A
FILING DATE: 31-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)751-6849
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid

STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-690-011A-53

Query Match 44.3%; Score 39; DB 2; Length 34;
Best Local Similarity 53.3%; Pred. No. 21;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 3 OKAAELNOKSKELEQ 17
.:|||:|:||||:
Db 20 KEAELEQENAELEQ 34

RESULT 10

US-08-690-011A-57
Sequence 57, Application US/08690011A
Patent No. 5942433
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
TITLE OF INVENTION: CELLULAR PROTEIN
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,011A
FILING DATE: 31-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)751-6849
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-690-011A-57

Query Match 44.3%; Score 39; DB 2; Length 34;
Best Local Similarity 53.3%; Pred. No. 21;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 3 OKAAELNOKSKELEQ 17
.:|||:|:||||:
Db 20 KEAELEQENAELEQ 34

RESULT 11

US-09-299-495F-53

Sequence 53, Application US/09299495F
Patent No. 6361968
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS WORD 97
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,495F
FILING DATE: 26-Apr-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: 08/690,011
FILING DATE: 31-JULY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-09-299-495F-53
Query Match 44.3%; Score 39; DB 3; Length 34;
Best Local Similarity 53.3%; Pred. No. 21;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Qy 3 OKAAELNOKSELEQ 17
Db 20 KEAELEQENAELEQ 34
RESULT 12
US-09-299-495F-57
Sequence 57, Application US/09299495F
Patent No. 6361968
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS WORD 97
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,495F
FILING DATE: 26-Apr-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: 08/690,011
FILING DATE: 31-JULY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 57:
US-09-299-495F-57
Query Match 44.3%; Score 39; DB 3; Length 34;
Best Local Similarity 53.3%; Pred. No. 21;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Qy 3 OKAAELNOKSELEQ 17
Db 20 KEAELEQENAELEQ 34
RESULT 13
US-09-621-976-5189
Sequence 5189, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 5189
LENGTH: 82
TYPE: PRT
ORGANISM: Homo sapiens
US-09-621-976-5189
Query Match 44.3%; Score 39; DB 4; Length 82;
Best Local Similarity 50.0%; Pred. No. 53;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 1 PFOKAAELNOKSELE 16
Db 31 PIVKVSRIYAKSKELD 46
RESULT 14

US-08-690-011A-19
Sequence 19, Application US/08690011A
Patent No. 5942433
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN
TITLE OF INVENTION: CELLULAR PROTEIN
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,011A
FILING DATE: 31-JUL-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)751-6849
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-690-011A-19
Query Match 44.3%; Score 39; DB 2; Length 98;
Best Local Similarity 50.0%; Pred. No. 64;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 3 QKAEINQSKKELEEQ 18
DB 42 KKAELQENALEEAE 57
RESULT 15
US-08-690-011A-3
Sequence 3, Application US/08690011A
Patent No. 5942433
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN
TITLE OF INVENTION: CELLULAR PROTEIN
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA

ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,011A
FILING DATE: 31-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)751-6849
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-690-011A-3
Query Match 43.2%; Score 38; DB 2; Length 26;
Best Local Similarity 43.8%; Pred. No. 22;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
QY 3 QKAEINQSKKELEEQ 18
DB 6 KKAELQENALEEAE 21
RESULT 16
US-09-299-495F-3
Sequence 3, Application US/09299495F
Patent No. 6361968
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS WORD 97
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,495F
FILING DATE: 26-Apr-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: 08/690,011
FILING DATE: 31-JULY-1996

ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-299-495F-3

Query Match 43.2%; Score 38; DB 3; Length 26;
Best Local Similarity 43.8%; Pred. No. 22;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 3 OKAAELNOKSKELEQ 18
Db 6 KEAELEQENAELEKE 21

RESULT 17
US-08-690-011A-38
Sequence 38, Application US/08690011A
Patent No. 5942433
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,011A
FILING DATE: 31-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-08-690-011A-38

Query Match 43.2%; Score 38; DB 2; Length 33;
Best Local Similarity 43.8%; Pred. No. 29;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 3 OKAAELNOKSKELEQ 18
Db 6 QRAEELARENELEKE 21

RESULT 18
US-09-299-495F-38
Sequence 38, Application US/09299495F
Patent No. 6361968
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS WORD 97
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,495F
FILING DATE: 26-Apr-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: 08/690,011
FILING DATE: 31-JULY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-299-495F-38

Query Match 43.2%; Score 38; DB 3; Length 33;
Best Local Similarity 43.8%; Pred. No. 29;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 3 OKAAELNOKSKELEQ 18
Db 6 QRAEELARENELEKE 21

RESULT 19
US-08-690-011A-54
Sequence 54, Application US/08690011A
Patent No. 5942433

GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,011A
FILING DATE: 31-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)751-6800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-690-011A-54

Query Match 43.2%; Score 38; DB 2; Length 34;
Best Local Similarity 43.8%; Pred. No. 29;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSELEQ 18
Db 6 QRAEELARENELEKE 21

RESULT 20
US-09-299-495F-54
Sequence 54, Application US/09299495F
Patent No. 6361968
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS WORD 97
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,495F
FILING DATE: 26-Apr-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: 08/690,011
FILING DATE: 31-JULY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-09-299-495F-54

Query Match 43.2%; Score 38; DB 3; Length 34;
Best Local Similarity 43.8%; Pred. No. 29;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 3 QKAAELNOKSELEQ 18
Db 6 QRAEELARENELEKE 21

RESULT 21
US-09-299-495F-11
Sequence 11, Application US/09299495F
Patent No. 6361968
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS WORD 97
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,495F
FILING DATE: 26-Apr-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: 08/690,011
FILING DATE: 31-JULY-1996
ATTORNEY/AGENT INFORMATION:


```

? NAME: Serunian, Leslie A.
? REGISTRATION NUMBER: 35,353
? REFERENCE/DOCKET NUMBER: 2026-4199US2
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212)758-4800
? TELEFAX: (212)751-6849
? INFORMATION FOR SEQ ID NO: 11:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 84 amino acids
? TYPE: amino acid
? STRANDEDNESS: unknown
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-299-495F-11

```

Query Match	43.2%	Score 38;	DB 3;	length 84;
Best Local Similarity	43.8%	Pred. No. 78;		
Matches	7;	Conservative	7;	Mismatches 2;
				Indels 0;
				Gaps 0;
QY	3	OKAAFLNOKSKTELEEQ	18	
DB	28	QAAELARENELEKE	43	

```

RESULT 22
US-09-134-000C-4168
: Sequence 4168, Application US/09134000C
: Patent No. 6617156
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
: FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
: PRIORITY OF INVENTION: 032796-032
: CURRENT APPLICATION NUMBER: US/09/134,000C
: CURRENT FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/055,778
: PRIOR FILING DATE: 1997-08-15
: NUMBER OF SEQ ID NOS: 6812
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4168
:
: LENGTH: 85
: TYPE: FRT
: ORGANISM: Enterococcus faecalis
US-09-134-000C-4168

```

	Query Match	43.2%	Score 38;	DB 4;	length 85;
	Best Local Similarity	56.2%	Pred. No. 79;		
	Matches	9; conservative	1; mismatches	6; indels	0; gaps
OY	2 TQKAELNOKSRELFQ	17			
Db	41 TDRIELNQVIRELER	56			

RESULT 23
 US-09-239-495F-17
 : Sequence 17, Application US/09299495F
 : Patent No. 6361868
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: VINSON, Charles R.
 : KRYLOV, Dmitry
 :
 : TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
 : INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
 : CELLULAR PROTEIN
 :
 : NUMBER OF SEQUENCES: 64
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
 : STREET: 345 Park Avenue
 : CITY: New York
 : STATE: NY
 : COUNTRY: USA
 : ZIP: 10154-0053
 :

```

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS WORD 97
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,495F
FILING DATE: 26-Apr-1999

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: 08/690,011
FILING DATE: 31-JULY-1996

ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-419US22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849

INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 86 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-299-495F-17

```

```

Query Match          43.2%  Score 38;  DB 3;  Length 86;
Best Local Similarity 43.8%  Pred. No. 80;
Matches             7;  Conservative 7;  Mismatches 2;  Indels 0;  Gaps 0;

OY      3  OKAAELNOKSEKELEEQ 18
      | : | | | : : : : | : :
Db      18  QRAEELARENNEEELEKE 33

```

RESULT 24
US-09-299-495F-52
Sequence 52. Application US/09299495F
Patent No. 6361968
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
KRYLOV, Dmitry
TITLE OF INVENTION: INTERACTION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & PINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS WORD 97
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,495F
FILING DATE: 26-Apr-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: 08/690,011

```

      FILING DATE: 31-JULY-1996
ATTORNEY/AGENT INFORMATION:
      NAME: Serunian, Leslie A.
      REGISTRATION NUMBER: 35,353
      REFERENCE/DOCKET NUMBER: 2026-4199US2
TELECOMMUNICATION INFORMATION:
      TELEPHONE: (212)758-4800
      TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 52:
      SEQUENCE CHARACTERISTICS:
          LENGTH: 86 amino acids
          TYPE: amino acid
          STRANDEDNESS: unknown
          TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-09-299-495F-52

Query Match           43.2% Score 38; DB 3; Length 86;
Best Local Similarity 43.8%; Pred. No. 80;
Matches    7; Conservative   7; Mismatches    2; Indels    0; Gaps    0;

QY              3 QKAEELNQSKELEQQ 18
                |||:::|||:
Db             18 QRAEELARENELEKE 33

RESULT 25
US-08-690-011A-11
Sequence 11, Application US/0890011A
Patent No. 5942433
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
TITLE OF INVENTION: CELLULAR PROTEIN
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSER: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690, 011A
FILING DATE: 31-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 88 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide

```

```

US-08-690-011A-11
Query Match      43.2%; Score 38; DB 2; Length 88;
Best Local Similarity 43.8%; Pred. No. 82;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY      3 QKAELNOKSELEEQ 18
      |||:::||||:
      28 QRAELARENELAKE 43

Db

RESULT 26
US-09-299-495F-19
; Sequence 19, Application US/09299495F
; Patent No. 6361968
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; KRYLOV, Dmitry
; TITLE OF INVENTION:
; EXTENSION OF A PROTEIN-PROTEIN
; INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; CELLULAR PROTEIN
;
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS WORD 97
; SOFTWARE: ASCII
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,495F
; FILING DATE: 26-Apr-1999
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; APPLICATION NUMBER: 08/690,011
; FILING DATE: 31-JULY-1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)758-4800
; TELEFAX: (212)751-6849
;
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-299-495F-19

Query Match      43.2%; Score 38; DB 3; Length 97;
Best Local Similarity 43.8%; Pred. No. 91;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY      3 QKAELNOKSELEEQ 18
      |||:::||||:
      28 QRAELARENELAKE 43

Db

RESULT 27
US-08-956-307B-2
; Sequence 2, Application US/08956307B

```

Patent No. 6090911
GENERAL INFORMATION:
APPLICANT: Petka, Wendy A.
APPLICANT: Tirrell, David A.
APPLICANT: Kevin P. McGrath
TITLE OF INVENTION: REVERSIBLE HYDROGELS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,307B
FILING DATE: 22-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07880/033001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-956-307B-2

Query Match 42.0%; Score 37; DB 3; Length 42;
Best Local Similarity 53.3%; Pred. No. 53;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KAAELNQSKELEEQ 18
Db 21 KAAELKQEVSRLENE 35

RESULT 28
US-09-015-030-7
Sequence 7, Application US/09015030
Patent No. 6303317
GENERAL INFORMATION:
APPLICANT: Albert, Thomas C.
APPLICANT: Sharma, Victoria A.
APPLICANT: Nautiyal, Shivan
TITLE OF INVENTION: Peptide Probes and Methods for Making the Same
FILE REFERENCE: A-64988-1/RFT/TAL
CURRENT APPLICATION NUMBER: US/09/015,030
CURRENT FILING DATE: 1998-01-28
PRIOR APPLICATION NUMBER: 60/036,219
PRIOR FILING DATE: 1997-01-28
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 43
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic.
US-09-015-030-7

Query Match 42.0%; Score 37; DB 3; Length 43;
Best Local Similarity 37.5%; Pred. No. 54;
Matches 6; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 3 QKAAELNQSKELEEQ 18
Db 14 EETKLEKKKDLBQK 29

RESULT 29
US-09-898-554-31
Sequence 31, Application US/09898554
Patent No. 6756228
GENERAL INFORMATION:
APPLICANT: TALU, ALAN R
APPLICANT: WELCH, CARIE L
APPLICANT: LIANG, CHIEH-PING
TITLE OF INVENTION: ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 1 (ATHSQ1) AND ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 2 (ATHSQ2)
FILE REFERENCE: 0575/64077
CURRENT APPLICATION NUMBER: US/09/898,554
CURRENT FILING DATE: 2001-07-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Version 3.1
SEQ ID NO 31
LENGTH: 46
TYPE: PRT
ORGANISM: Murinae gen. sp.
FEATURE:
NAME/KEY: MISC FEATURE
OTHER INFORMATION: ISOFORM 1 REPEAT #3
US-09-898-554-31

Query Match 42.0%; Score 37; DB 4; Length 46;
Best Local Similarity 63.6%; Pred. No. 58;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 ELNQSKELEEQ 17
Db 16 KLNESKQOE 26

RESULT 30
US-09-898-554-36
Sequence 36, Application US/09898554
Patent No. 6756228
GENERAL INFORMATION:
APPLICANT: TALL, ALAN R
APPLICANT: WELCH, CARIE L
APPLICANT: LIANG, CHIEH-PING
TITLE OF INVENTION: ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 1 (ATHSQ1) AND ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 2 (ATHSQ2)
FILE REFERENCE: 0575/64077
CURRENT APPLICATION NUMBER: US/09/898,554
CURRENT FILING DATE: 2001-07-02
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Version 3.1
SEQ ID NO 36
LENGTH: 46
TYPE: PRT
ORGANISM: Murinae gen. sp.
FEATURE:
NAME/KEY: MISC FEATURE
OTHER INFORMATION: ISOFORM 7 REPEAT#3
US-09-898-554-36

Query Match 42.0%; Score 37; DB 4; Length 46;
Best Local Similarity 63.6%; Pred. No. 58;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 ELNQSKELEEQ 17
Db 16 KLNESKQOE 26

```
Db      16 KLNKSKKEQEE 26

RESULT 31
US-09-898-554-37
; Sequence 37, Application US/09898554
; Patent No. 6756228
; GENERAL INFORMATION:
; APPLICANT: TAIL, ALAN R
; APPLICANT: WELCH, CARRIE L
; APPLICANT: LIANG, CHEN-PING
; TITLE OF INVENTION: ATHEROSCLEROSIS SUSCEPTIBILITY GENE LOCUS 1 (ATHSQ1) AND ATHEROS
; FILE REFERENCE: 0575/64077
; CURRENT APPLICATION NUMBER: US/09/898,554
; PRIOR FILING DATE: 2001-07-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
; FEATURE:
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: ISOFORM 8 REPEAT#3
US-09-898-554-37

Query Match      42.0%; Score 37; DB 4; Length 46;
Best Local Similarity 63.6%; Pred. No. 58;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      7 ELNKSKELEQ 17
       :|:|:|:|:|:|:
Db      16 KLNKSKKEQEE 26

RESULT 32
US-09-513-999C-5756
; Sequence 5756, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59. US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5756
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-5756

Query Match      42.0%; Score 37; DB 4; Length 52;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      3 OKAAELNKSKELEQ 18
       :|:|:|:|:|:|:
Db      16 KKAALKAQRIKELQEE 31

RESULT 33
US-09-583-110-4322
; Sequence 4322, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4322
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4322

Query Match      42.0%; Score 37; DB 4; Length 67;
Best Local Similarity 53.8%; Pred. No. 87;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      6 AELNKSKELEQ 18
       ||:|:|:|:|:|:
Db      38 AEIAQKREIEQR 50

RESULT 34
US-09-490-291-11
; Sequence 11, Application US/09490291
; Patent No. 6620917
; GENERAL INFORMATION:
; APPLICANT: Mello, Charlene M.
; APPLICANT: Arcidiacono, Steven
; TITLE OF INVENTION: No. 6620917el Purification and Fiber Spinning Techniques for
; FILE REFERENCE: ARMY-03665
; CURRENT APPLICATION NUMBER: US/09/490,291
; CURRENT FILING DATE: 2000-01-20
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Nephila clavipes
US-09-490-291-11

Query Match      42.0%; Score 37; DB 4; Length 74;
Best Local Similarity 53.3%; Pred. No. 97;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      4 KAAELNKSKELEQ 18
       |||||:|:|:|:
Db      35 KAAELKQEVRLNE 49

RESULT 35
US-08-936-165A-343
; Sequence 343, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
```

TITLE OF INVENTION: Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 343:
SEQUENCE CHARACTERISTICS:
LENGTH: 82 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-936-165A-343

Query Match 42.0%; Score 37; DB 3; Length 82;
Best Local Similarity 46.7%; Pred. No. 1.1e+02;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 OKAEINOKSKELQ 17
Db 54 KKAELDEKKKKLKE 68

RESULT 36
US-08-452-592B-7
Sequence 7, Application US/08452592B
Patent No. 5712366
GENERAL INFORMATION:
APPLICANT: McGrath, Kevin P.
APPLICANT: Kaplan, David L.
TITLE OF INVENTION: Fabrication of Nanoscale Materials Using
TITLE OF INVENTION: Self-Assembling Proteins
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: U.S. Army Soldier Systems Command
STREET: AMSCC-CC (Patent Counsel)
CITY: Natick
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 01760-5035
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: DOS 6.22
SOFTWARE: WordPerfect Version 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,592B
FILING DATE: 25 May 1995
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/068,948
FILING DATE: 25 May 1993
ATTORNEY/AGENT INFORMATION:
NAME: Donahue, Richard J.
NAME: Lamming, John H.
NAME: Ranucci, Vincent J.
REGISTRATION NUMBER: 22,062 (Donahue)
REGISTRATION NUMBER: 34,857 (Lamming)
REGISTRATION NUMBER: 29,579 (Ranucci)
REFERENCE/DOCKET NUMBER: NA-1096D
TELEPHONE: 508-233-4510
TELEFAX: 508-233-5167
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-452-592B-7

Query Match 42.0%; Score 37; DB 1; Length 84;
Best Local Similarity 53.3%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 4 KKAELNOKSKELQ 18
Db 21 KKAELNOKSKELQ 35

RESULT 37
US-08-452-592B-8
Sequence 8, Application US/08452592B
Patent No. 5712366
GENERAL INFORMATION:
APPLICANT: McGrath, Kevin P.
APPLICANT: Kaplan, David L.
TITLE OF INVENTION: Fabrication of Nanoscale Materials Using
TITLE OF INVENTION: Self-Assembling Proteins
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: U.S. Army Soldier Systems Command
STREET: AMSCC-CC (Patent Counsel)
CITY: Natick
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 01760-5035
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: DOS 6.22
SOFTWARE: WordPerfect Version 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,592B
FILING DATE: 25 May 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/068,948
FILING DATE: 25 May 1993
ATTORNEY/AGENT INFORMATION:
NAME: Donahue, Richard J.
NAME: Lamming, John H.
NAME: Ranucci, Vincent J.
REGISTRATION NUMBER: 22,062 (Donahue)
REGISTRATION NUMBER: 34,857 (Lamming)
REGISTRATION NUMBER: 29,579 (Ranucci)
REFERENCE/DOCKET NUMBER: NA-1096D
TELEPHONE: 508-233-4510
TELEFAX: 508-233-5167
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids

```
;
;      TYPE: amino acid
;      TOPOLOGY: linear
;
US-08-452-592B-8
;
Query Match      42.0%; Score 37; DB 1; Length 84;
Best Local Similarity 53.3%; Pred. No. 1.1e+02;
Matches      8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      4 KAAELNOKSELEEQ 18
      |||||:|:|:
Db      21 KAAELKQEVSRLENE 35

RESULT 38
US-08-452-592B-9
; Sequence 9, Application US/08452592B
; Patent No. 5712366
; GENERAL INFORMATION:
; APPLICANT: McGrath, Kevin P.
; APPLICANT: Kaplan, David L.
; TITLE OF INVENTION: Fabrication of Nanoscale Materials Using
; TITLE OF INVENTION: Self-Assembling Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: U.S. Army Soldier Systems Command
; STREET: AMSCC-CC (Patent Counsel)
; CITY: Natick
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01760-5035
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
; OPERATING SYSTEM: DOS 6.22
; SOFTWARE: WordPerfect Version 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/452,592B
; FILING DATE: 25 May 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/068,948
; FILING DATE: 25 May 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Donahue, Richard J.
; NAME: Lamming, John H.
; NAME: Ranucci, Vincent J.
; REGISTRATION NUMBER: 22,062 (Donahue)
; REGISTRATION NUMBER: 34,857 (Lamming)
; REGISTRATION NUMBER: 29,579 (Ranucci)
; REFERENCE/DOCKET NUMBER: NA-1096D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 508-233-4510
; TELEFAX: 508-233-5167
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-452-592B-9
;
Query Match      42.0%; Score 37; DB 1; Length 84;
Best Local Similarity 53.3%; Pred. No. 1.1e+02;
Matches      8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      4 KAAELNOKSELEEQ 18
      |||||:|:|:
Db      21 KAAELKQEVSRLENE 35

RESULT 39
US-09-513-999C-5991
; Sequence 5991, Application US/09513999C
; Patent No. 6783961
;
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59,US2,REG
; PATENT NO. 6783961
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5991
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-513-999C-5991
;
Query Match      42.0%; Score 37; DB 4; Length 87;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches      8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      3 QKAAELNOKSELE 16
      :|||:|||||
Db      59 RKVAECQKLEKELE 72

RESULT 40
US-09-107-532A-5271
; Sequence 5271, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; CORRESPONDENCE ADDRESS: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 893-5007
; TELEFAX: (781) 893-8277
; INFORMATION FOR SEQ ID NO: 5271:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
```

FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...98
SEQUENCE DESCRIPTION: SEQ ID NO: 5271:
US-09-107-532A-5271

Query Match 42.0%; Score 37; DB 4; Length 98;
Best Local Similarity 63.6%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 8 LNOKSKELEQ 18
Db 37 LNOKMKOLEDQ 47

RESULT 41
US-09-902-540-12287
Sequence 12287, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(115849)B
CURRENT APPLICATION NUMBER: US/09/902,540
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 12287
LENGTH: 97
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-12287

Query Match 41.5%; Score 36.5; DB 4; Length 97;
Best Local Similarity 69.2%; Pred. No. 1.5e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Oy 1 PT-OKAELNOKS 12
Db 31 PTPOKAQLBROKA 43

RESULT 42
US-08-690-011A-2
Sequence 2, Application US/08690011A
Patent No. 5942433
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,011A
FILING DATE: 31-JUL-1996
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-690-011A-2

Query Match 40.9%; Score 36; DB 2; Length 25;
Best Local Similarity 46.7%; Pred. No. 43;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Oy 3 OKAELNOKSKELEQ 17
Db 6 KEAELEQENAELE 20

RESULT 43
US-09-299-495F-2
Sequence 2, Application US/09299495F
Patent No. 6361968
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: MS WORD 97
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,495F
FILING DATE: 26-Apr-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: 08/690,011
FILING DATE: 31-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)751-6849
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids

```

; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-299-495F-2

```

```

Query Match 40.9%; Score 36; DB 3; Length 25;
Best Local Similarity 46.7%; Pred. No. 43;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 3 QKAALNQSKKELEQ 17
::|||::|||:
Db 6 KEAELEQENAELE 20

```

```

RESULT 44
US-09-640-211A-622
; Sequence 622, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 622
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-640-211A-622

```

```

Query Match 40.9%; Score 36; DB 4; Length 79;
Best Local Similarity 58.3%; Pred. No. 1.5e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 5 AAELNQSKKELE 16
:|:|:|||||
Db 62 SVEVKNKSKKELE 73

```

```

RESULT 45
US-09-640-211A-2181
; Sequence 2181, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2181
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-640-211A-2181

```

```

Query Match 40.9%; Score 36; DB 4; Length 83;
Best Local Similarity 58.3%; Pred. No. 1.6e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 5 AAELNQSKKELE 16
:|:|:|||||
Db 66 SVEVKNKSKKELE 77

```

```

Search completed: August 29, 2005, 17:14:21
Job time : 44 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2005, 16:38:49 ; Search time 164 Seconds
(without alignments)
42.449 Million cell updates/sec

Title: US-10-031-289-1331
Perfect score: 88
Sequence: 1 PTQKAHLNKSKELEQQ 18

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 1171057

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp1980s:*\n2: geneseqp1990s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	50.0	43	2	AAW69983 Yeast tro
2	43	48.9	94	2	AAR21561 dnag carb
3	42	47.7	25	3	AAW08381 Peptide u
4	42	47.7	35	3	AAW08365 Amino aci
5	42	47.7	35	3	AAW08367 Amino aci
6	42	47.7	35	3	AAW08345 Amino aci
7	42	47.7	35	3	AAW08363 Amino aci
8	42	47.7	35	3	AAW08384 Peptide u
9	42	47.7	35	6	AAW00697 Protein m
10	42	47.7	35	6	AAW00670 Protein m
11	42	47.7	35	6	AAW00691 Protein m
12	42	47.7	35	6	AAW00693 Protein m
13	42	47.7	35	6	AAW00672 Protein m
14	42	47.7	35	6	AAW00653 Protein m
15	42	47.7	35	6	AAW00695 Protein m
16	42	47.7	35	6	AAW00686 Protein m
17	42	47.7	35	6	AAW00687 Protein m
18	42	47.7	35	6	AAW00674 Protein m
19	42	47.7	43	3	AAW08383 Peptide u
20	42	47.7	43	3	AAW08364 Amino aci
21	42	47.7	43	6	AAW00694 Protein m
22	42	47.7	43	6	AAW00680 Protein m
23	42	47.7	43	6	AAW00671 Protein m
24	41	46.6	94	2	AAW00956 4Heptad-P
25	40	45.5	43	3	AAW08375 Imobilis

26	40	45.5	43	6	ADA00682 Protein m
27	40	45.5	82	3	AAW01789 Human sec
28	39.5	44.9	46	6	AAW79886 LOX-1 iso
29	39.5	44.9	46	6	AAW79891 LOX-1 iso
30	39.5	44.9	46	6	AAW79889 LOX-1 iso
31	39	44.3	27	3	AAW37026 Bcl2 poly
32	39	44.3	31	3	AAW08379 Peptide u
33	39	44.3	31	3	AAW08395 Peptide u
34	39	44.3	31	6	AAW09302 Peptide 4
35	39	44.3	33	6	AAW00689 Protein m
36	39	44.3	40	8	AAW97089 Botulinum
37	39	44.3	43	3	AAW08344 Amino aci
38	39	44.3	43	6	AAW00652 Protein m
39	39	44.3	97	2	AAW00949 CWV500-4H
40	38.5	43.8	83	8	AAW17053 Nanoarcha
41	38	43.2	25	3	AAW08382 Peptide u
42	38	43.2	26	2	AAW00940 2Heptadma
43	38	43.2	31	6	AAW08387 Peptide u
44	38	43.2	31	6	AAW00698 Protein m
45	38	43.2	33	2	AAW00961 Heptad pe
46	38	43.2	35	4	AAW89483 Binding p
47	38	43.2	35	4	AAW00681 Protein m
48	38	43.2	35	6	AAW00683 Protein m
49	38	43.2	35	6	AAW00679 Protein m
50	38	43.2	73	6	AAW1006 Human rev
51	38	43.2	77	6	AAW1006 Human rev
52	38	43.2	79	7	AAW10609 Structura
53	38	43.2	79	7	AAW10609 Structura
54	38	43.2	79	8	AAW15628 Nucleatin
55	38	43.2	83	8	AAW15624 Nucleatin
56	38	43.2	84	2	AAW00945 CWV500-4H
57	38	43.2	84	5	AAW33504 Human ORF
58	38	43.2	85	7	AAW10608 Structura
59	38	43.2	85	7	AAW10608 Structura
60	38	43.2	85	7	AAW10608 Structura
61	38	43.2	85	8	AAW15627 Nucleatin
62	38	43.2	87	2	AAW00948 4Heptadfo
63	38	43.2	91	5	AAW03791 Human ORF
64	38	43.2	95	2	AAW76095 Mouse ost
65	37	42.0	42	3	AAW10558 Colled co
66	37	42.0	43	2	AAW69984 Yeast tro
67	37	42.0	46	6	AAW79894 LOX-1 iso
68	37	42.0	46	6	AAW79888 LOX-1 iso
69	37	42.0	46	6	AAW79893 LOX-1 iso
70	37	42.0	46	8	AAW80874 Mouse dna
71	37	42.0	52	3	AAW01675 Human sec
72	37	42.0	57	4	AAW21279 Peptide #
73	37	42.0	57	4	AAW43612 Peptide #
74	37	42.0	57	4	AAW37506 Peptide #
75	37	42.0	57	4	AAW26566 Protein #
76	37	42.0	57	4	AAW77355 Human bon
77	37	42.0	57	4	AAW6551 Human bra
78	37	42.0	57	4	AAW65988 Human liv
79	37	42.0	57	5	AAW46371 Human pep
80	37	42.0	61	8	AAW05314 Human gen
81	37	42.0	62	3	AAW85685 Streptoco
82	37	42.0	67	6	AAW02712 S. pneumo
83	37	42.0	67	8	AAW47807 Streptoco
84	37	42.0	74	4	AAW82613 Recognin
85	37	42.0	80	7	AAW10607 Structura
86	37	42.0	80	8	AAW15626 Nucleatin
87	37	42.0	84	2	AAW77583 Stephyloc
88	37	42.0	84	2	AAW43025 Artificia
89	37	42.0	84	2	AAW43027 Artificia
90	37	42.0	84	2	AAW43026 Artificia
91	37	42.0	87	3	AAW01910 Human sec
92	37	42.0	93	6	AAW71064 Human adi
93	37	42.0	94	4	AAW78767 Human per
94	37	42.0	95	2	AAW36943 Protein e
95	37	42.0	98	7	AAW36944 E. faeciu
96	36	40.9	23	3	AAW08388 Peptide u
97	36	40.9	23	6	AAW00699 Protein m
98	36	40.9	25	2	AAW00939 2Heptadma

99	36	40.9	33	3	AAb21682	Colied co
100	36	40.9	33	4	AAg62757	Amino aci

ALIGNMENTS

RESULT 1
AAW69983
ID AAW69983 standard; peptide; 43 AA.
XX
XX AAW69983;
AC
XX
XX
DT 20-OCT-1998 (first entry)
XX
DE Yeast tropomyosin (TPM1) coiled coil target sequence.
XX
XX Yeast: tropomyosin; TPM1; peptide probe; coil protein; biological probe;
KW staining; target protein; affinity purification; coprecipitation;
KM oligomerisation; colon cancer.
XX
XX Saccharomyces cerevisiae.
OS
XX
XX W09832848-A1.
FN
XX
XX 30-JUL-1998.
PD
XX
XX 28-JAN-1998; 98WO-US001871.
PF
XX
XX 28-JAN-1997; 97US-0036219P.
PR
XX
XX (REGC) UNIV CALIFORNIA.
PA
XX
XX Alber TC, Allen V, Nautiyal S;
PI
XX
XX WPI; 1998-427944/36.
DR
XX
XX
PT Heterospecific polypeptide probes with coiled coil regions - useful for
PT detecting and staining target proteins, tagging cells, screening for
PT mutant protein forms and diagnosing colon cancer.
XX
XX
PS Example 6; Page 59; 84pp; English.

Query Match	50.0%;	Score 44;	DB 2;	Length 43;
Best Local Similarity	50.0%;	Pred. No. 21;		
Matches	8;	Conservative	5;	Mismatches 3;
				Indels 0;
				Gaps 0;

QY 3 QKAE LNQSKSELEQQ 18
:| | :| :| :| :

D_b 14 EKYEEJKEKNKDLQE 29

RESULT 2
AAR21561
ID AAR21561 standard; peptide; 94 AA.

DT	25-MAR-2003	(revised)
DT	04-JUN-1992	(first entry)

DE dnag carboxy terminus.

KW Macromolecular system (MMS) operon; initiation of DNA, RNA and protein synthesis; primase; DNA replication.

OS *Listeria monocytogenes*.

PN EP472434-A.

PD 26-FEB-1992.

22-AUG-1991; 91EP-00307750.

PR 23-AUG-1990; 90US-00572191.

PA (BAYU) BAYLOR COLLEGE MEDICINE.

PA (ABBO) ABBOTT LAB.

PI Lupski JR, Katz L;

DR WPI; 1992-066623/09.

DR N-PSDB; AAQ21651.

PT Anti:sense oligo:

are used to treat.

PS Disclosure; Page 12; 48pp; English.

CC The dnaG C-terminal protein sequence was deduced from the MMS operon DNA

including a protein-protein interaction with the primosome complex, a

polymerase activity to synthesise the primer RNA and plays a role in the

See also AAR23002, 9 and AAR21562, 3. (Updated on 25-MAR-2003 to correct PA

cc
cc
cc

cc
cc
cc

Sequence 94 AA;

Query Match	48.9%	Score 43	DB 2	Length 94
Best Local Similarity	60.0%	Pred. No. 70		
Matches 9; Conservative		2; Mismatches	4; Indels	0; Gaps

QY	4	KAEI.NÖKSKELEÖQ	18
Db	74	KRFKEÖKKKELEÖQ	88

RESULT 3	
AAB08381	
ID	AAB08381 standard; peptide; 25 AA.

AC AAB08381;

DT 20-DEC-2000 (first entry)

Peptide used to assay for peptidase activity.

Binding partner; protein modification; post-translational modification; modulator; coiled-coil structure.

XX Unidentified.
OS
XX WO200050902-A2.
PN
XX 31-AUG-2000.
PD
XX 25-FEB-2000; 2000WO-GB000669.
XX
XX 25-FEB-1999; 99GB-00004398.
PR
XX (FLUO-) FLUORESCENCE LTD.
XX
XX Colyer J, Craig RK, Maschio A, Mezna M;
PI WPI; 2000-572119/53.
DR
XX High throughput assay for monitoring modification of polypeptides and
PT modulation of the modifications.
XX
XX Example 8; Page 79; 128pp; English.
PS
XX The specification describes a method for analysing a sample. The method
CC comprises immobilising a polypeptide to a physical support, contacting
CC the immobilised polypeptide with a test sample which may contain an agent
CC capable of modifying the immobilised polypeptide, contacting the
CC immobilised polypeptide with a binding partner polypeptide, where
CC association of both polypeptide is dependent on the modification state of
CC the immobilised polypeptide, and measuring the association of the binding
CC partner polypeptide to the immobilised polypeptide. The polypeptides,
CC support and methods can be used to analyse a sample to determine if
CC modification of a polypeptide is taking place and to identify modulators
CC of the modification. This is useful for monitoring the post-translational
CC modification of proteins. AAB08379-82 represent peptides, which are used
CC in the method of the invention, to assay for amino- and carboxypeptidase
CC activity
XX
XX Sequence 25 AA;
SQ

Query Match 47.7%; Score 42; DB 3; Length 25;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Qy 3 OKAAELNOKSKKELEQQ 18
Db 6 QKIAQLKQKNAQLKQK 21

RESULT 4
AAB08365
ID AAB08365 standard; peptide; 35 AA.
XX
XX AAB08365;
AC
XX 20-DEC-2000 (first entry)
DT
XX Amino acid sequence of a coiled-coil peptide.
DE
XX Binding partner; protein modification; post-translational modification;
KW modulator; coiled-coil structure.
XX
XX Unidentified.
OS
XX Key Location/Qualifiers
FH Misc-difference 30 /note= "label attachment site"
FT
XX
XX WO200050902-A2.
PN
XX 31-AUG-2000.
PD
XX 25-FEB-2000; 2000WO-GB000669.
PF
XX

PR 25-FEB-1999; 99GB-00004398.
XX
XX (FLUO-) FLUORESCENCE LTD.
PA
XX Colyer J, Craig RK, Maschio A, Mezna M;
PI WPI; 2000-572119/53.
DR
XX High throughput assay for monitoring modification of polypeptides and
PT modulation of the modifications.
XX
XX Disclosure; Page 61; 128pp; English.
PS
XX The specification describes a method for analysing a sample. The method
CC comprises immobilising a polypeptide to a physical support, contacting
CC the immobilised polypeptide with a test sample which may contain an agent
CC capable of modifying the immobilised polypeptide, contacting the
CC immobilised polypeptide with a binding partner polypeptide, where
CC association of both polypeptide is dependent on the modification state of
CC the immobilised polypeptide, and measuring the association of the binding
CC partner polypeptide to the immobilised polypeptide. The polypeptides,
CC support and methods can be used to analyse a sample to determine if
CC modification of a polypeptide is taking place and to identify modulators
CC of the modification. This is useful for monitoring the post-translational
CC modification of proteins. AAB08364-65 represent binding partners which
CC have coiled-coil structures, and may be used in the method of the
CC invention to assay for thrombin activity
XX
XX Sequence 35 AA;
SQ

Query Match 47.7%; Score 42; DB 3; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Qy 3 OKAAELNOKSKKELEQQ 18
Db 6 QKIAQLKQKNAQLKQK 21

RESULT 5
AAB08367
ID AAB08367 standard; peptide; 35 AA.
XX
XX AAB08367;
AC
XX 20-DEC-2000 (first entry)
DT
XX Amino acid sequence of a coiled-coil peptide.
DE
XX Binding partner; protein modification; post-translational modification;
KW modulator; coiled-coil structure.
XX
XX Unidentified.
OS
XX Key Location/Qualifiers
FH Misc-difference 30 /note= "label attachment site"
FT
XX
XX WO200050902-A2.
PN
XX 31-AUG-2000.
PD
XX 25-FEB-2000; 2000WO-GB000669.
PF
XX 25-FEB-1999; 99GB-00004398.
PR
XX (FLUO-) FLUORESCENCE LTD.
PA
XX Colyer J, Craig RK, Maschio A, Mezna M;
PI WPI; 2000-572119/53.
DR
XX High throughput assay for monitoring modification of polypeptides and
PT

PT modulation of the modifications.
XX
PS Disclosure; Page 62; 128pp; English.
XX
CC The specification describes a method for analysing a sample. The method
CC comprises immobilising a polypeptide to a physical support, contacting
CC the immobilised polypeptide with a test sample which may contain an agent
CC capable of modifying the immobilised polypeptide, contacting the
CC immobilised polypeptide with a binding partner polypeptide, where
CC association of both polypeptide is dependent on the modification state of
CC the immobilised polypeptide, and measuring the association of the binding
CC partner polypeptide to the immobilised polypeptide. The polypeptides,
CC support and methods can be used to analyse a sample to determine if
CC modification of a polypeptide is useful for monitoring the post-translational
CC of the modification. This is useful for monitoring the post-translational
CC modification of proteins. AAB08362-67 represent binding partners which
CC have coiled-coil structures, and may be used in the method of the
CC invention to assay for tobacco etch virus (TEV) protease activity
XX
SQ Sequence 35 AA;
XX
Query Match 47.7%; Score 42; DB 3; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 3 OKAAELNQSKKELEQQ 18
DB 6 QKTAQLKQKNAQLKQK 21
XX
RESULT 6
AAB08345
ID AAB08345 standard; peptide; 35 AA.
AC AAB08345;
XX
DT 20-DEC-2000 (first entry)
XX
DE Amino acid sequence of a coiled-coil structure partner.
XX
KW Binding partner; protein modification; post-translational modification;
KW modulator; coiled-coil structure.
XX
OS Unidentified.
XX
PN WO200050902-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-GB000669.
XX
PR 25-FEB-1999; 99GB-00004398.
XX
PA (FLUO-) FLUORESCENCE LTD.
XX
PI Colyer J, Craig RK, Maschio A, Mezna M;
XX WPI; 2000-572119/53.
XX
DR WPI; 2000-572119/53.
XX
PT High throughput assay for monitoring modification of polypeptides and
XX modulation of the modifications.
XX
PS Disclosure; Page 21; 128pp; English.
XX
CC The specification describes a method for analysing a sample. The method
CC comprises immobilising a polypeptide to a physical support, contacting
CC the immobilised polypeptide with a test sample which may contain an agent
CC capable of modifying the immobilised polypeptide, contacting the
CC immobilised polypeptide with a binding partner polypeptide, where
CC association of both polypeptide is dependent on the modification state of
CC the immobilised polypeptide, and measuring the association of the binding
CC partner polypeptide to the immobilised polypeptide. The polypeptides,
CC support and methods can be used to analyse a sample to determine if

CC modification of a polypeptide is taking place and to identify modulators
CC of the modification. This is useful for monitoring the post-translational
CC modification of proteins. AAB08344-45 represent binding partners which
CC have coiled-coil structures, and may be used in the method of the
CC invention
XX
SQ Sequence 35 AA;
XX
Query Match 47.7%; Score 42; DB 3; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 3 OKAAELNQSKKELEQQ 18
DB 6 QKTAQLKQKNAQLKQK 21
XX
RESULT 7
AAB08363
ID AAB08363 standard; peptide; 35 AA.
XX
AC AAB08363;
XX
DT 20-DEC-2000 (first entry)
XX
DE Amino acid sequence of a coiled-coil peptide.
XX
KW Binding partner; protein modification; post-translational modification;
KW modulator; coiled-coil structure.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 30 /note="label attachment site"
FT
XX
PN WO200050902-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-GB000669.
XX
PR 25-FEB-1999; 99GB-00004398.
XX
PA (FLUO-) FLUORESCENCE LTD.
XX
PI Colyer J, Craig RK, Maschio A, Mezna M;
XX WPI; 2000-572119/53.
XX
DR WPI; 2000-572119/53.
XX
PT High throughput assay for monitoring modification of polypeptides and
XX modulation of the modifications.
XX
PS Disclosure; Page 61; 128pp; English.
XX
CC The specification describes a method for analysing a sample. The method
CC comprises immobilising a polypeptide to a physical support, contacting
CC the immobilised polypeptide with a test sample which may contain an agent
CC capable of modifying the immobilised polypeptide, contacting the
CC immobilised polypeptide with a binding partner polypeptide, where
CC association of both polypeptide is dependent on the modification state of
CC the immobilised polypeptide, and measuring the association of the binding
CC partner polypeptide to the immobilised polypeptide. The polypeptides,
CC support and methods can be used to analyse a sample to determine if
CC modification of a polypeptide is taking place and to identify modulators
CC of the modification. This is useful for monitoring the post-translational
CC modification of proteins. AAB08362-63 represent binding partners which
CC have coiled-coil structures, and may be used in the method of the
CC invention to assay for chymotrypsin activity
XX
SQ Sequence 35 AA;
XX
Query Match 47.7%; Score 42; DB 3; Length 35;

Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 OKAAELNOKSKLEEQ 18
Db 6 OKIAQLKOKNAQLKOK 21

RESULT 8
AAB08384
ID AAB08384 standard; peptide; 35 AA.

XX AAB08384;

XX 20-DEC-2000 (first entry)

XX Peptide used to assay for tobacco etch virus protease activity.

XX Binding partner; protein modification; post-translational modification;

KM modulator; coiled-coil structure.

XX Unidentified.

XX W0200050902-A2.

XX 31-AUG-2000.

XX 25-FEB-2000; 2000WO-GB000669.

XX 25-FEB-1999; 99GB-00004398.

XX (FLUO-) FLUORESCENCE LTD.

XX Colyer J, Craig RK, Maschio A, Mezna M;

XX WPI; 2000-572119/53.

XX High throughput assay for monitoring modification of polypeptides and

PT modulation of the modifications.

XX Example 9; Page 80; 128pp; English.

CC The specification describes a method for analysing a sample. The method
CC comprises immobilising a polypeptide to a physical support, contacting
CC the immobilised polypeptide with a test sample which may contain an agent
CC capable of modifying the immobilised polypeptide, contacting the
CC immobilised polypeptide with a binding partner polypeptide, where
CC association of both polypeptide is dependent on the modification state of
CC the immobilised polypeptide, and measuring the association of the binding
CC partner polypeptide to the immobilised polypeptide. The polypeptides,
CC support and methods can be used to analyse a sample to determine if
CC modification of a polypeptide is taking place and to identify modulators
CC of the modification. This is useful for monitoring the post-translational
CC modification of proteins. AAB08383-84 represent binding partners which
CC are used, in the method of the invention, to assay for tobacco etch virus
CC (TEV) protease activity

XX Sequence 35 AA;

XX Query Match 47.7%; Score 42; DB 3; Length 35;
XX Best Local Similarity 50.0%; Pred. No. 34;
XX Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 OKAAELNOKSKLEEQ 18
Db 6 OKIAQLKOKNAQLKOK 21

RESULT 9
ADA00697
ID ADA00697 standard; peptide; 35 AA.
XX ADA00697;
AC

XX 06-NOV-2003 (first entry)

DE Protein modification analysis method associated partner peptide #13.

XX protein modification analysis method; protein modifying agent;

KM surface plasma resonance; scintillation proximity assay; proteolysis;

KM phosphorylation; acylation; glycosylation; farnesylation; geranylation;

KM ubiquitination; prenylation; sentrinisation;

KM adenosine diphosphate ribosylation; ADP ribosylation;

KM protein modification; human.

XX Homo sapiens.

XX US2003032054-A1.

XX 13-FEB-2003.

XX 26-FEB-1999; 99US-00259658.

XX 26-FEB-1999; 99US-00259658.

XX (WILL) WILLIAMS K M.

XX Colyer J, Craig RK, Maschio A, Mezna M;

XX WPI; 2003-605702/57.

XX Analysis method, useful for detecting or monitoring the activity of a
XX modulator of a polypeptide modifying agent, comprises immobilizing a
XX polypeptide to a physical support, contacting with another polypeptide
XX and assaying.

XX Example 11; Page 29; 32pp; English.

CC The invention describes an analysis method, useful for detecting or
CC monitoring the activity of a modulator of a polypeptide modifying agent.
CC The method comprises: providing a polypeptide pair, where association of
CC the polypeptides is detectable and modification of at least one
CC polypeptide results in modulation of the association; contacting
CC immobilised first polypeptide with the second polypeptide; and assaying
CC the modification. The immobilised polypeptide and/or the binding partner
CC polypeptide is/are associated with a label such as a fluorescent label or
CC a radioactive label. The detectable signal is generated by an interaction
CC between the labels comprising energy transfer. The association is
CC measured by monitoring the molecular mass of the hybrid species
CC comprising the second polypeptide associated with the first polypeptide,
CC by surface plasma resonance, by scintillation proximity assay, or by
CC using an antibody. Assaying the modification comprises assaying
CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
CC geranylation, ubiquitination, prenylation, sentrinisation, adenosine
CC diphosphate (ADP)-ribosylation, or the reversal of any of these
CC modifications. The method is used for analysing a sample, i.e. detecting
CC or monitoring the activity of a modulator of a polypeptide modifying
CC agent. The invention enables only one polypeptide to be susceptible to a
CC single modification, allowing only one modification event per complex,
CC thus resulting in increased output of the assay, as effectively every
CC modification event which takes place will have an effect on the readout.
CC This is the amino acid sequence of a partner peptide that binds to and
CC alters the modification of a template peptide ADA00652. The invention
CC provides a method for monitoring the modification state of the template
CC peptide on binding of its partner.

XX Sequence 35 AA;

XX Query Match 47.7%; Score 42; DB 6; Length 35;
XX Best Local Similarity 50.0%; Pred. No. 34;
XX Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 OKAAELNOKSKLEEQ 18
Db 6 OKIAQLKOKNAQLKOK 21

RESULT 10
 ID ADA00670 standard; peptide; 35 AA.
 AC
 XX ADA00670;
 XX
 DT 06-NOV-2003 (first entry)
 DE Protein modification analysis method associated partner peptide #2.
 XX
 KW protein modification analysis method; protein modifying agent;
 KW surface plasma resonance; scintillation proximity assay; proteolysis;
 KW phosphorylation; acylation; glycosylation; farnesylation; geranylation;
 KW ubiquitination; prenylation; sentrinisation;
 KW adenosine diphosphate ribosylation; ADP ribosylation;
 KW protein modification; human.
 XX
 OS Homo sapiens.
 XX
 PN US2003032054-A1.
 PD
 XX 13-FEB-2003.
 XX
 XX 26-FEB-1999; 99US-00259658.
 XX
 XX 26-FEB-1999; 99US-00259658.
 XX
 XX (WILL/) WILLIAMS K M.
 XX
 PI Colyer J, Craig RK, Maschio A, Mezna M;
 XX WPI; 2003-605702/57.
 DR
 XX
 XX Analysis method, useful for detecting or monitoring the activity of a
 PT modulator of a polypeptide modifying agent, comprises immobilizing a
 PT polypeptide to a physical support, contacting with another polypeptide
 PT and assaying.
 PT
 XX
 PS Disclosure; Page 21, 32pp; English.
 XX
 CC The invention describes an analysis method, useful for detecting or
 CC monitoring the activity of a modulator of a polypeptide modifying agent.
 CC The method comprises: providing a polypeptide pair, where association of
 CC the polypeptides is detectable and modification of at least one
 CC polypeptide results in modulation of the association; contacting
 CC immobilised first polypeptide with the second polypeptide; and assaying
 CC the modification. The immobilised polypeptide and/or the binding partner
 CC polypeptide is/are associated with a label such as a fluorescent label or
 CC a radioactive label. The detectable signal is generated by an interaction
 CC between the labels comprising energy transfer. The association is
 CC measured by monitoring the molecular mass of the hybrid species
 CC comprising the second polypeptide associated with the first polypeptide,
 CC by surface plasma resonance, by scintillation proximity assay, or by
 CC using an antibody. Assaying the modification comprises assaying
 CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
 CC geranylation, ubiquitination, prenylation, sentrinisation, adenosine
 CC diphosphate (ADP)-ribosylation, or the reversal of any of these
 CC modifications. The method is used for analysing a sample, i.e. detecting
 CC or monitoring the activity of a modulator of a polypeptide modifying
 CC agent. The invention enables only one polypeptide to be susceptible to a
 CC single modification, allowing only one modification event per complex,
 CC thus resulting in increased output of the assay, as effectively every
 CC modification event which takes place will have an effect on the readout.
 CC This is the amino acid sequence of a partner peptide that binds to and
 CC alters the modification of a template peptide ADA00652. The invention
 CC provides a method for monitoring the modification state of the template
 CC peptide on binding of it's partner.
 CC
 XX
 SQ Sequence 35 AA;
 Query Match 47.7%; Score 42; DB 6; Length 35;
 Best Local Similarity 50.0%; Pred. No. 34;

	Matches	8;	Conservative	5;	Mismatches	3;	Indels	0;	Gaps	0;
Qy	3	OKAEINOKSKKEEQ	18							
Db	6	OKTAQLKQKNAQLKQK	21							

RESULT 11
 ID ADA00691 standard; peptide; 35 AA.
 AC
 XX ADA00691;
 XX
 DT 06-NOV-2003 (first entry)
 DE Protein modification analysis method associated partner peptide #10.
 XX
 KW protein modification analysis method; protein modifying agent;
 KW surface plasma resonance; scintillation proximity assay; proteolysis;
 KW phosphorylation; acylation; glycosylation; farnesylation; geranylation;
 KW ubiquitination; prenylation; sentrinisation;
 KW adenosine diphosphate ribosylation; ADP ribosylation;
 KW protein modification; human.
 XX
 OS Homo sapiens.
 XX
 PN US2003032054-A1.
 PD
 XX 13-FEB-2003.
 XX
 XX 26-FEB-1999; 99US-00259658.
 XX
 XX 26-FEB-1999; 99US-00259658.
 XX
 XX (WILL/) WILLIAMS K M.
 XX
 PI Colyer J, Craig RK, Maschio A, Mezna M;
 XX WPI; 2003-605702/57.
 DR
 XX
 XX Analysis method, useful for detecting or monitoring the activity of a
 PT modulator of a polypeptide modifying agent, comprises immobilizing a
 PT polypeptide to a physical support, contacting with another polypeptide
 PT and assaying.
 PT
 XX
 PS Example 7; Page 27, 32pp; English.
 XX
 CC The invention describes an analysis method, useful for detecting or
 CC monitoring the activity of a modulator of a polypeptide modifying agent.
 CC The method comprises: providing a polypeptide pair, where association of
 CC the polypeptides is detectable and modification of at least one
 CC polypeptide results in modulation of the association; contacting
 CC immobilised first polypeptide with the second polypeptide; and assaying
 CC the modification. The immobilised polypeptide and/or the binding partner
 CC polypeptide is/are associated with a label such as a fluorescent label or
 CC a radioactive label. The detectable signal is generated by an interaction
 CC between the labels comprising energy transfer. The association is
 CC measured by monitoring the molecular mass of the hybrid species
 CC comprising the second polypeptide associated with the first polypeptide,
 CC by surface plasma resonance, by scintillation proximity assay, or by
 CC using an antibody. Assaying the modification comprises assaying
 CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
 CC geranylation, ubiquitination, prenylation, sentrinisation, adenosine
 CC diphosphate (ADP)-ribosylation, or the reversal of any of these
 CC modifications. The method is used for analysing a sample, i.e. detecting
 CC or monitoring the activity of a modulator of a polypeptide modifying
 CC agent. The invention enables only one polypeptide to be susceptible to a
 CC single modification, allowing only one modification event per complex,
 CC thus resulting in increased output of the assay, as effectively every
 CC modification event which takes place will have an effect on the readout.
 CC This is the amino acid sequence of a partner peptide that binds to and
 CC alters the modification of a template peptide ADA00652. The invention
 CC provides a method for monitoring the modification state of the template

CC peptide on binding of it's partner.
XX
SQ Sequence 35 AA;
Query Match 47.7%; Score 42; DB 6; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
OY 3 OKAAELNOKSKLEEQ 18
DB 6 QKIAQLKQKNAQLKQK 21
RESULT 12
IDA00693 standard; peptide; 35 AA.
XX ADA00693;
XX 06-NOV-2003 (first entry)
XX Protein modification analysis method associated partner peptide #11.
XX protein modification analysis method; protein modifying agent;
KW surface plasma resonance; scintillation proximity assay; proteolysis;
KW phosphorylation; acylation; glycosylation; farnesylation; geranylation;
KW ubiquitination; prenylation; sentrinisation;
KW adenosine diphosphate ribosylation; ADP ribosylation;
KW protein modification; tobacco etch virus.
XX OS Tobacco etch virus.
XX US2003032054-A1.
XX 13-FEB-2003.
XX 26-FEB-1999; 99US-00259658.
XX PF 26-FEB-1999; 99US-00259658.
XX PR 26-FEB-1999; 99US-00259658.
XX PA (WILL/) WILLIAMS K M.
XX PI Colyer J, Craig RK, Maschio A, Mezna M;
XX WPI; 2003-605702/57.
XX Analysis method, useful for detecting or monitoring the activity of a
PT modulator of a polypeptide modifying agent, comprises immobilizing a
PT polypeptide to a physical support, contacting with another polypeptide
PT and assaying.
XX
XX Example 9; Page 29; 32pp; English.
XX The invention describes an analysis method, useful for detecting or
CC monitoring the activity of a modulator of a polypeptide modifying agent.
CC The method comprises: providing a polypeptide pair, where association of
CC the polypeptides is detectable and modification of at least one
CC polypeptide results in modulation of the association; contacting
CC immobilised first polypeptide with the second polypeptide; and assaying
CC the modification. The immobilised polypeptide and/or the binding partner
CC polypeptide is/are associated with a label such as a fluorescent label or
CC a radioactive label. The detectable signal is generated by an interaction
CC between the labels comprising energy transfer. The association is
CC measured by monitoring the molecular mass of the hybrid species
CC comprising the second polypeptide associated with the first polypeptide,
CC by surface plasma resonance, by scintillation proximity assay, or by
CC using an antibody. Assaying the modification comprises assaying
CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
CC geranylation, ubiquitination, prenylation, sentrinisation, adenosine
CC diphosphate (ADP)-ribosylation, or the reversal of any of these
CC modifications. The method is used for analysing a sample, i.e. detecting
CC or monitoring the activity of a modulator of a polypeptide modifying
CC agent. The invention enables only one polypeptide to be susceptible to a

CC single modification, allowing only one modification event per complex,
CC thus resulting in increased output of the assay, as effectively every
CC modification event takes place will have an effect on the readout.
CC This is the amino acid sequence of a partner peptide that binds to and
CC alters the modification of a template peptide ADA00652. The invention
CC provides a method for monitoring the modification state of the template
CC peptide on binding of it's partner.
XX
XX SQ Sequence 35 AA;
Query Match 47.7%; Score 42; DB 6; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
OY 3 OKAAELNOKSKLEEQ 18
DB 6 QKIAQLKQKNAQLKQK 21
RESULT 13
IDA00672 standard; peptide; 35 AA.
XX ADA00672;
XX 06-NOV-2003 (first entry)
XX Protein modification analysis method associated partner peptide #3.
XX protein modification analysis method; protein modifying agent;
KW surface plasma resonance; scintillation proximity assay; proteolysis;
KW phosphorylation; acylation; glycosylation; farnesylation; geranylation;
KW ubiquitination; prenylation; sentrinisation;
KW adenosine diphosphate ribosylation; ADP ribosylation;
KW protein modification; human.
XX OS Homo sapiens.
XX US2003032054-A1.
XX 13-FEB-2003.
XX 26-FEB-1999; 99US-00259658.
XX PF 26-FEB-1999; 99US-00259658.
XX PR 26-FEB-1999; 99US-00259658.
XX PA (WILL/) WILLIAMS K M.
XX PI Colyer J, Craig RK, Maschio A, Mezna M;
XX WPI; 2003-605702/57.
XX Analysis method, useful for detecting or monitoring the activity of a
PT modulator of a polypeptide modifying agent, comprises immobilizing a
PT polypeptide to a physical support, contacting with another polypeptide
PT and assaying.
XX
XX Disclosure; Page 22; 32pp; English.
XX The invention describes an analysis method, useful for detecting or
CC monitoring the activity of a modulator of a polypeptide modifying agent.
CC The method comprises: providing a polypeptide pair, where association of
CC the polypeptides is detectable and modification of at least one
CC polypeptide results in modulation of the association; contacting
CC immobilised first polypeptide with the second polypeptide; and assaying
CC the modification. The immobilised polypeptide and/or the binding partner
CC polypeptide is/are associated with a label such as a fluorescent label or
CC a radioactive label. The detectable signal is generated by an interaction
CC between the labels comprising energy transfer. The association is
CC measured by monitoring the molecular mass of the hybrid species
CC comprising the second polypeptide associated with the first polypeptide,
CC by surface plasma resonance, by scintillation proximity assay, or by
CC using an antibody. Assaying the modification comprises assaying

CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation, geranyl-
CC acylation, ubiquitination, prenylation, serratination, adenosine
CC dihydrophate (ADP)-riboseylation, or the reversal of any of these
CC modifications. The method is used for analysing a sample, i.e. detecting
CC or monitoring the activity of a modulator of a polypeptide modifying
CC agent. The invention enables only one polypeptide to be susceptible to a
CC single modification, allowing only one modification event per complex,
CC thus resulting in increased output of the assay, as effectively every
CC modification event which takes place will have an effect on the readout.
CC This is the amino acid sequence of a partner peptide that binds to and
CC alters the modification of a template peptide ADA00652. The invention
CC provides a method for monitoring the modification state of the template
CC peptide on binding of its partner.
XX
SQ Sequence 35 AA;
Query Match 47.7%; Score 42; DB 6; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Db 3 OKAAELNOKSKELREQ 18
6 QKTAQLKQKNAQLKQK 21
RESULT 14
ADA00653
ID ADA00653 standard; peptide; 35 AA.
XX
AC ADA00653;
XX
DT 06-NOV-2003 (first entry)
XX
DE Protein modification analysis method associated partner peptide #1.
XX
KW protein modification analysis method; protein modifying agent;
KW surface plasma resonance; scintillation proximity assay; proteolysis;
KW phosphorylation; acylation; glycosylation; farnesylation; geranyl-
KW ubiquitination; prenylation; serratination;
KW adenosine diphosphate ribosylation; ADP ribosylation;
KW protein modification.
XX
OS Unidentified.
XX
PN US2003032054-A1.
XX
PD 13-FEB-2003.
XX
PF 26-FEB-1999; 99US-00259658.
XX
PR 26-FEB-1999; 99US-00259658.
XX
PA (WILL/) WILLIAMS K M.
XX
PI Colyer J, Craig RK, Maschio A, Mezna M;
XX
DR WPI; 2003-605702/57.
XX
PT Analysis method, useful for detecting or monitoring the activity of a
PT modulator of a polypeptide modifying agent, comprises immobilizing a
PT polypeptide to a physical support, contacting with another polypeptide
PT and assaying.
XX
PS Disclosure; Page 8; 32pp; English.
XX
CC The invention describes an analysis method, useful for detecting or
CC monitoring the activity of a modulator of a polypeptide modifying agent.
CC The method comprises: providing a polypeptide pair, where association of
CC the polypeptides is detectable and modification of at least one
CC polypeptide results in modulation of the association; contacting
CC immobilised first polypeptide with the second polypeptide; and assaying
CC the modification. The immobilised polypeptide and/or the binding partner
CC polypeptide is/are associated with a label such as a fluorescent label or

CC a radioactive label. The detectable signal is generated by an interaction
CC between the labels comprising energy transfer. The association is
CC measured by monitoring the molecular mass of the hybrid species
CC comprising the second polypeptide associated with the first polypeptide,
CC by surface plasma resonance, by scintillation proximity assay, or by
CC using an antibody. Assaying the modification comprises assaying
CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
CC geranyl-acylation, ubiquitination, prenylation, serratination, adenosine
CC dihydrophate (ADP)-riboseylation, or the reversal of any of these
CC modifications. The method is used for analysing a sample, i.e. detecting
CC or monitoring the activity of a modulator of a polypeptide modifying
CC agent. The invention enables only one polypeptide to be susceptible to a
CC single modification, allowing only one modification event per complex,
CC thus resulting in increased output of the assay, as effectively every
CC modification event which takes place will have an effect on the readout.
CC This is the amino acid sequence of a partner peptide that binds to and
CC alters the modification of a template peptide ADA00652. The invention
CC provides a method for monitoring the modification state of the template
CC peptide on binding of its partner.
XX
SQ Sequence 35 AA;
Query Match 47.7%; Score 42; DB 6; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Db 3 OKAAELNOKSKELREQ 18
6 QKTAQLKQKNAQLKQK 21
RESULT 15
ADA00695
ID ADA00695 standard; peptide; 35 AA.
XX
AC ADA00695;
XX
DT 06-NOV-2003 (first entry)
XX
DE Protein modification analysis method associated partner peptide #12.
XX
KW protein modification analysis method; protein modifying agent;
KW surface plasma resonance; scintillation proximity assay; proteolysis;
KW phosphorylation; acylation; glycosylation; farnesylation; geranyl-
KW ubiquitination; prenylation; serratination;
KW adenosine diphosphate ribosylation; ADP ribosylation;
KW protein modification; human.
XX
OS Homo sapiens.
XX
PN US2003032054-A1.
XX
PD 13-FEB-2003.
XX
PF 26-FEB-1999; 99US-00259658.
XX
PR 26-FEB-1999; 99US-00259658.
XX
PA (WILL/) WILLIAMS K M.
XX
PI Colyer J, Craig RK, Maschio A, Mezna M;
XX
DR WPI; 2003-605702/57.
XX
PT Analysis method, useful for detecting or monitoring the activity of a
PT modulator of a polypeptide modifying agent, comprises immobilizing a
PT polypeptide to a physical support, contacting with another polypeptide
PT and assaying.
XX
PS Example 10; Page 29; 32pp; English.
XX
CC The invention describes an analysis method, useful for detecting or
CC monitoring the activity of a modulator of a polypeptide modifying agent.

CC The method comprises: providing a polypeptide pair, where association of
 CC the polypeptides is detectable and modification of at least one
 CC polypeptide results in modulation of the association; contacting
 CC immobilised first polypeptide with the second polypeptide; and assaying
 CC the modification. The immobilised polypeptide and/or the binding partner
 CC polypeptide is/are associated with a label such as a fluorescent label or
 CC a radioactive label. The detectable signal is generated by an interaction
 CC between the labels comprising energy transfer. The association is
 CC measured by monitoring the molecular mass of the hybrid species
 CC comprising the second polypeptide associated with the first polypeptide,
 CC by surface plasma resonance, by scintillation proximity assay, or by
 CC using an antibody. Assaying the modification comprises assaying
 CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
 CC geranylation, ubiquitination, prenylation, sentrinisation, adenosine
 CC diphosphate (ADP)-ribosylation, or the reversal of any of these
 CC modifications. The method is used for analysing a sample, i.e. detecting
 CC or monitoring the activity of a modulator of a polypeptide modifying
 CC agent. The invention enables only one polypeptide to be susceptible to a
 CC single modification, allowing only one modification event per complex,
 CC thus resulting in increased output of the assay, as effectively every
 CC modification event which takes place will have an effect on the readout.
 CC This is the amino acid sequence of a partner peptide that binds to and
 CC alters the modification of a template peptide ADA00652. The invention
 CC provides a method for monitoring the modification state of the template
 CC peptide on binding of it's partner.

XX Sequence 35 AA;

Query Match 47.7%; Score 42; DB 6; Length 35;
 Best Local Similarity 50.0%; Pred. No. 34;
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 OKAAELNOKSKELFEQ 18
 |||:|:|:|:|:|:
 Db 6 QKIAQLKQKNAQLKQK 21

RESULT 16
 ADA00686

ID ADA00686 standard; peptide; 35 AA.

XX ADA00686;

DT 06-NOV-2003 (first entry)

XX Protein modification analysis method associated partner peptide #8.

XX protein modification analysis method; protein modifying agent;
 KW surface plasma resonance; scintillation proximity assay; proteolysis;
 KW phosphorylation; acylation; glycosylation; farnesylation; geranylation;
 KW ubiquitination; prenylation; sentrinisation;
 KW adenosine diphosphate ribosylation; ADP ribosylation;
 KW protein modification.

XX Unidentified.

PN US2003032054-A1.

PD 13-FEB-2003.

PF 26-FEB-1999; 99US-00259658.

PR 26-FEB-1999; 99US-00259658.

PA (WILL/) WILLIAMS K M.

PI Colyer J, Craig RK, Maschio A, Mezna M;

DR WPI; 2003-605702/57.

XX Analysis method, useful for detecting or monitoring the activity of a
 PT modulator of a polypeptide modifying agent, comprises immobilizing a
 PT polypeptide to a physical support, contacting with another polypeptide

PT and assaying.

PS Example 6; Page 27; 32pp; English.

XX The invention describes an analysis method, useful for detecting or
 CC monitoring the activity of a modulator of a polypeptide modifying agent.
 CC The method comprises: providing a polypeptide pair, where association of
 CC the polypeptides is detectable and modification of at least one
 CC polypeptide results in modulation of the association; contacting
 CC immobilised first polypeptide with the second polypeptide; and assaying
 CC the modification. The immobilised polypeptide and/or the binding partner
 CC polypeptide is/are associated with a label such as a fluorescent label or
 CC a radioactive label. The detectable signal is generated by an interaction
 CC between the labels comprising energy transfer. The association is
 CC measured by monitoring the molecular mass of the hybrid species
 CC comprising the second polypeptide associated with the first polypeptide,
 CC by surface plasma resonance, by scintillation proximity assay, or by
 CC using an antibody. Assaying the modification comprises assaying
 CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
 CC geranylation, ubiquitination, prenylation, sentrinisation, adenosine
 CC diphosphate (ADP)-ribosylation, or the reversal of any of these
 CC modifications. The method is used for analysing a sample, i.e. detecting
 CC or monitoring the activity of a modulator of a polypeptide modifying
 CC agent. The invention enables only one polypeptide to be susceptible to a
 CC single modification, allowing only one modification event per complex,
 CC thus resulting in increased output of the assay, as effectively every
 CC modification event which takes place will have an effect on the readout.
 CC This is the amino acid sequence of a partner peptide that binds to and
 CC alters the modification of a template peptide ADA00652. The invention
 CC provides a method for monitoring the modification state of the template
 CC peptide on binding of it's partner.

SO Sequence 35 AA;

Query Match 47.7%; Score 42; DB 6; Length 35;
 Best Local Similarity 50.0%; Pred. No. 34;
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 OKAAELNOKSKELFEQ 18
 |||:|:|:|:|:|:
 Db 6 QKIAQLKQKNAQLKQK 21

RESULT 17
 ADA00687

ID ADA00687 standard; peptide; 35 AA.

XX ADA00687;

DT 06-NOV-2003 (first entry)

XX Protein modification analysis method associated template peptide #28.

XX protein modification analysis method; protein modifying agent;
 KW surface plasma resonance; scintillation proximity assay; proteolysis;
 KW phosphorylation; acylation; glycosylation; farnesylation; geranylation;
 KW ubiquitination; prenylation; sentrinisation;
 KW adenosine diphosphate ribosylation; ADP ribosylation;
 KW protein modification; yeast; human.

XX Homo sapiens.

PN US2003032054-A1.

PD 13-FEB-2003.

PF 26-FEB-1999; 99US-00259658.

PR 26-FEB-1999; 99US-00259658.

PA (WILL/) WILLIAMS K M.

PI Colyer J, Craig RK, Maschio A, Mezna M;

XX WPI; 2003-605702/57.
DR Analysis method, useful for detecting or monitoring the activity of a
PT modulator of a polypeptide modifying agent, comprises immobilizing a
PT polypeptide to a physical support, contacting with another polypeptide
PT and assaying.
XX
PS Example 7, Page 27, 32pp; English.
XX
XX The invention describes an analysis method, useful for detecting or
CC monitoring the activity of a modulator of a polypeptide modifying agent.
CC The method comprises: providing a polypeptide pair, where association of
CC the polypeptides is detectable and modification of at least one
CC polypeptide results in modulation of the association; contacting
CC immobilised first polypeptide with the second polypeptide; and assaying
CC the modification. The immobilised polypeptide and/or the binding partner
CC polypeptide is/are associated with a label such as a fluorescent label or
CC a radioactive label. The detectable signal is generated by an interaction
CC between the labels comprising energy transfer. The association is
CC measured by monitoring the molecular mass of the hybrid species
CC comprising the second polypeptide associated with the first polypeptide,
CC by surface plasma resonance, by scintillation proximity assay, or by
CC using an antibody. Assaying the modification comprises assaying
CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
CC geranylolation, ubiquitination, prenylation, serratination, adenosine
CC diophosphate (ADP)-ribosylation, or the reversal of any of these
CC modifications. The method is used for analysing a sample, i.e. detecting
CC or monitoring the activity of a modulator of a polypeptide modifying
CC agent. The invention enables only one polypeptide to be susceptible to a
CC single modification, allowing only one modification event per complex,
CC thus resulting in increased output of the assay, as effectively every
CC modification event which takes place will have an effect on the readout.
CC This is the amino acid sequence of a template peptide containing
CC modification sites affected by the binding of a partner peptide ADA0653.
CC The invention provides a method for monitoring the modification state of
CC the template peptide on binding of its partner.
XX
SQ Sequence 35 AA;
Query Match 47.7%; Score 42; DB 6; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 3 OKAAELNQSKELEQQ 18
DB 6 QKIAQLKQKNQALQK 21
RESULT 18
ADA00674
ID ADA00674 standard; peptide; 35 AA.
XX
XX ADA00674;
DT 06-NOV-2003 (first entry)
XX
XX Protein modification analysis method associated partner peptide #4.
XX
XX protein modification analysis method; protein modifying agent;
KW surface plasma resonance; scintillation proximity assay; proteolysis;
KW phosphorylation; acylation; glycosylation; farnesylation; geranylolation;
KW ubiquitination; prenylation; serratination;
KW adenosine diophosphate ribosylation; ADP ribosylation;
KW protein modification; tobacco etch virus protease; tobacco etch virus.
XX
XX Tobacco etch virus.
XX
XX OS US2003032054-A1.
XX
XX 13-FEB-2003.
XX
XX 26-FEB-1999; 99US-00259658.

XX
PR 26-FEB-1999; 99US-00259658.
XX
XX (WILL) WILLIAMS K M.
XX
XX Colyer J, Craig RK, Maschio A, Mezna M;
XX
XX WPI; 2003-605702/57.
XX
XX
XX Analysis method, useful for detecting or monitoring the activity of a
PT modulator of a polypeptide modifying agent, comprises immobilizing a
PT polypeptide to a physical support, contacting with another polypeptide
PT and assaying.
XX
PS Disclosure; Page 22, 32pp; English.
XX
XX The invention describes an analysis method, useful for detecting or
CC monitoring the activity of a modulator of a polypeptide modifying agent.
CC The method comprises: providing a polypeptide pair, where association of
CC the polypeptides is detectable and modification of at least one
CC polypeptide results in modulation of the association; contacting
CC immobilised first polypeptide with the second polypeptide; and assaying
CC the modification. The immobilised polypeptide and/or the binding partner
CC polypeptide is/are associated with a label such as a fluorescent label or
CC a radioactive label. The detectable signal is generated by an interaction
CC between the labels comprising energy transfer. The association is
CC measured by monitoring the molecular mass of the hybrid species
CC comprising the second polypeptide associated with the first polypeptide,
CC by surface plasma resonance, by scintillation proximity assay, or by
CC using an antibody. Assaying the modification comprises assaying
CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
CC geranylolation, ubiquitination, prenylation, serratination, adenosine
CC diophosphate (ADP)-ribosylation, or the reversal of any of these
CC modifications. The method is used for analysing a sample, i.e. detecting
CC or monitoring the activity of a modulator of a polypeptide modifying
CC agent. The invention enables only one polypeptide to be susceptible to a
CC single modification, allowing only one modification event per complex,
CC thus resulting in increased output of the assay, as effectively every
CC modification event which takes place will have an effect on the readout.
CC This is the amino acid sequence of a partner peptide that binds to and
CC alters the modification of a template peptide ADA0652. The invention
CC provides a method for monitoring the modification state of the template
CC peptide on binding of its partner.
XX
SQ Sequence 35 AA;
Query Match 47.7%; Score 42; DB 6; Length 35;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 3 OKAAELNQSKELEQQ 18
DB 6 QKIAQLKQKNQALQK 21
RESULT 19
AAB08383
ID AAB08383 standard; peptide; 43 AA.
XX
XX AAB08383;
AC
XX 20-DEC-2000 (first entry)
DT
XX Peptide used to assay for tobacco etch virus protease activity.
DB
XX Binding partner; protein modification; post-translational modification;
KW modulator; coiled-coil structure.
XX
XX Unidentified.
XX
XX OS WO200005902-A2.
XX
XX 31-AUG-2000.

XX 25-FEB-2000; 2000WO-GB000669.
XX 25-FEB-1999; 99GB-00004398.
XX (FLUO-) FLUORESCENCE LTD.
XX
PI Colyer J, Craig RK, Maschio A, Mezna M;
XX WPI; 2000-572119/53.
XX
PT High throughput assay for monitoring modification of polypeptides and
XX modulation of the modifications.
XX
PS Example 9; Page 80; 128pp; English.
XX
CC The specification describes a method for analysing a sample. The method
CC comprises immobilising a polypeptide to a physical support, contacting
CC the immobilised polypeptide with a test sample which may contain an agent
CC capable of modifying the immobilised polypeptide, contacting the
CC immobilised polypeptide with a binding partner polypeptide, where
CC association of both polypeptide is dependent on the modification state of
CC the immobilised polypeptide, and measuring the association of the binding
CC partner polypeptide to the immobilised polypeptide. The polypeptides,
CC support and methods can be used to analyse a sample to determine if
CC modification of a polypeptide is taking place and to identify modulators
CC of the modification. This is useful for monitoring the post-translational
CC modification of proteins. AAB08364-64 represent binding partners which
CC are used, in the method of the invention, to assay for tobacco etch virus
CC (TEV) protease activity
XX
SQ Sequence 43 AA;
XX
Query Match 47.7%; Score 42; DB 3; Length 43;
Best Local Similarity 43.8%; Pred. No. 42;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
XX
QY 3 OKAAELNOKSKELKQ 18
DB 14 QETNQLKQENKRLKQ 29
XX
*
RESULT 20
AAB08364
ID AAB08364 standard; peptide; 43 AA.
XX
AC AAB08364;
XX
DT 20-DEC-2000 (first entry)
XX
DE Amino acid sequence of a coiled-coil peptide.
XX
XX Binding partner; protein modification; post-translational modification;
XX modulator; coiled-coil structure.
XX
OS Unidentified.
XX
XX Key Location/Qualifiers
FH Cleavage-site 24
FT /note= "thrombin cleavage site"
FT Misc-difference 39
FT /note= "fluorophore attachment site"
XX
XX WO200050902-A2.
XX
XX 31-AUG-2000.
XX
XX 25-FEB-2000; 2000WO-GB000669.
XX
XX 25-FEB-1999; 99GB-00004398.
XX
XX (FLUO-) FLUORESCENCE LTD.
XX

PI Colyer J, Craig RK, Maschio A, Mezna M;
XX WPI; 2000-572119/53.
XX
PT High throughput assay for monitoring modification of polypeptides and
XX modulation of the modifications.
XX
PS Disclosure; Page 61; 128pp; English.
XX
CC The specification describes a method for analysing a sample. The method
CC comprises immobilising a polypeptide to a physical support, contacting
CC the immobilised polypeptide with a test sample which may contain an agent
CC capable of modifying the immobilised polypeptide, contacting the
CC immobilised polypeptide with a binding partner polypeptide, where
CC association of both polypeptide is dependent on the modification state of
CC the immobilised polypeptide, and measuring the association of the binding
CC partner polypeptide to the immobilised polypeptide. The polypeptides,
CC support and methods can be used to analyse a sample to determine if
CC modification of a polypeptide is taking place and to identify modulators
CC of the modification. This is useful for monitoring the post-translational
CC modification of proteins. AAB08364-65 represent binding partners which
CC have coiled-coil structures, and may be used in the method of the
CC invention to assay for thrombin activity
XX
SQ Sequence 43 AA;
XX
Query Match 47.7%; Score 42; DB 3; Length 43;
Best Local Similarity 43.8%; Pred. No. 42;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
XX
QY 3 OKAAELNOKSKELKQ 18
DB 14 QETNQLKQENKRLKQ 29
XX
RESULT 21
ADA00694
ID ADA00694 standard; peptide; 43 AA.
XX
AC ADA00694;
XX
DT 06-NOV-2003 (first entry)
XX
DE Protein modification analysis method associated template peptide #32.
XX
XX protein modification analysis method; protein modifying agent;
XX surface plasma resonance; scintillation proximity assay; proteolysis;
XX phosphorylation; acylation; glycosylation; farnesylation; geranylation;
XX ubiquitination; prenylation; sentrinisation;
XX adenosine diphosphate ribosylation; ADP ribosylation;
XX protein modification; yeast; human.
XX
XX Homo sapiens.
XX
XX US2003032054-A1.
XX
XX 13-FEB-2003.
XX
XX 26-FEB-1999; 99US-00259658.
XX
XX 26-FEB-1999; 99US-00259658.
XX
XX (WILL) WILLIAMS K M.
XX
XX Colyer J, Craig RK, Maschio A, Mezna M;
XX WPI; 2003-605702/57.
XX
XX
PT Analysis method, useful for detecting or monitoring the activity of a
PT modulator of a polypeptide modifying agent, comprises immobilizing a
PT polypeptide to a physical support, contacting with another polypeptide
PT and assaying.
XX

PS Example 10; Page 29; 32pp; English.

XX The invention describes an analysis method, useful for detecting or
CC monitoring the activity of a modulator of a polypeptide modifying agent.
CC The method comprises: providing a polypeptide pair, where association of
CC the polypeptide is detectable and modification of at least one
CC polypeptide results in modulation of the association; contacting
CC immobilised first polypeptide with the second polypeptide; and assaying
CC the modification. The immobilised polypeptide and/or the binding partner
CC polypeptide is/are associated with a label such as a fluorescent label or
CC a radioactive label. The detectable signal is generated by an interaction
CC between the labels comprising energy transfer. The association is
CC measured by monitoring the molecular mass of the hybrid species
CC comprising the second polypeptide associated with the first polypeptide,
CC by surface plasma resonance, by scintillation proximity assay, or by
CC using an antibody. Assaying the modification comprises assaying
CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
CC geranylation, ubiquitination, prenylation, sentrinisation, adenosine
CC diphosphate (ADP)-ribosylation, or the reversal of any of these
CC modifications. The method is used for analysing a sample, i.e. detecting
CC or monitoring the activity of a modulator of a polypeptide modifying
CC agent. The invention enables only one modification event per complex,
CC thus resulting in increased output of the assay, as effectively every
CC modification event which takes place will have an effect on the readout.
CC This is the amino acid sequence of a template peptide containing
CC modification sites affected by the binding of a partner peptide ADA00653.
CC The invention provides a method for monitoring the modification state of
CC the template peptide on binding of its partner.

XX Sequence 43 AA;

Query Match 47.7%; Score 42; DB 6; Length 43;
Best Local Similarity 43.8%; Pred. No. 42;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 3 QKAELNQSKRELBOQ 18

Db 14 QETALQLEGNRLQEOE 29

RESULT 22

ADA00680
ID ADA00680 standard; peptide; 43 AA.

XX ADA00680;

XX 06-NOV-2003 (first entry)

DE Protein modification analysis method associated template peptide #24.

XX protein modification analysis method; protein modifying agent;
KW surface plasma resonance; scintillation proximity assay; proteolysis;
KW phosphorylation; acylation; glycosylation; farnesylation; geranylation;
KW ubiquitination; prenylation; sentrinisation;
KW adenosine diphosphate ribosylation; ADP ribosylation;
KW protein modification.

XX Unidentified.

XX US2003032054-A1.

PN 13-FEB-2003.

PD 26-FEB-1999; 99US-00259658.

PR 26-FEB-1999; 99US-00259658.

XX (WILL/) WILLIAMS K M.

PA Colyer J, Craig RK, Maschio A, Mezna M;

XX WPI; 2003-605702/57.

XX Analysis method, useful for detecting or monitoring the activity of a
PT modulator of a polypeptide modifying agent, comprises immobilizing a
PT polypeptide to a physical support, contacting with another polypeptide
PT and assaying.

PS Example 4; Page 26; 32pp; English.

XX The invention describes an analysis method, useful for detecting or
CC monitoring the activity of a modulator of a polypeptide modifying agent.
CC The method comprises: providing a polypeptide pair, where association of
CC the polypeptide is detectable and modification of at least one
CC polypeptide results in modulation of the association; contacting
CC immobilised first polypeptide with the second polypeptide; and assaying
CC the modification. The immobilised polypeptide and/or the binding partner
CC polypeptide is/are associated with a label such as a fluorescent label or
CC a radioactive label. The detectable signal is generated by an interaction
CC between the labels comprising energy transfer. The association is
CC measured by monitoring the molecular mass of the hybrid species
CC comprising the second polypeptide associated with the first polypeptide,
CC by surface plasma resonance, by scintillation proximity assay, or by
CC using an antibody. Assaying the modification comprises assaying
CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
CC geranylation, ubiquitination, prenylation, sentrinisation, adenosine
CC diphosphate (ADP)-ribosylation, or the reversal of any of these
CC modifications. The method is used for analysing a sample, i.e. detecting
CC or monitoring the activity of a modulator of a polypeptide modifying
CC agent. The invention enables only one modification event per complex,
CC thus resulting in increased output of the assay, as effectively every
CC modification event which takes place will have an effect on the readout.
CC This is the amino acid sequence of a template peptide containing
CC modification sites affected by the binding of a partner peptide ADA00653.
CC The invention provides a method for monitoring the modification state of
CC the template peptide on binding of its partner.

XX Sequence 43 AA;

Query Match 47.7%; Score 42; DB 6; Length 43;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 QKAELNQSKRELBOQ 18

Db 14 QETALQLEGNRLQEOE 29

RESULT 23

ADA00671
ID ADA00671 standard; peptide; 43 AA.

XX ADA00671;

XX 06-NOV-2003 (first entry)

DE Protein modification analysis method associated template peptide #18.

XX protein modification analysis method; protein modifying agent;
KW surface plasma resonance; scintillation proximity assay; proteolysis;
KW phosphorylation; acylation; glycosylation; farnesylation; geranylation;
KW ubiquitination; prenylation; sentrinisation;
KW adenosine diphosphate ribosylation; ADP ribosylation;
KW protein modification; human.

XX Unidentified.

XX US2003032054-A1.

PN 13-FEB-2003.

PD 26-FEB-1999; 99US-00259658.

PR 26-FEB-1999; 99US-00259658.

XX (WILL/) WILLIAMS K M.
 PA Colyer J, Craig RK, Maschio A, Mezna M;
 XX WPI; 2003-605702/57.
 DR
 XX
 PT Analysis method, useful for detecting or monitoring the activity of a
 PT modulator of a polypeptide modifying agent, comprises immobilising a
 PT polypeptide to a physical support, contacting with another polypeptide
 PT and assaying.
 XX
 PS Disclosure; Page 22; 32pp; English.
 XX
 CC The invention describes an analysis method, useful for detecting or
 CC monitoring the activity of a modulator of a polypeptide modifying agent.
 CC The method comprises: providing a polypeptide pair, where association of
 CC the polypeptides is detectable and modification of at least one
 CC polypeptide results in modulation of the association; contacting
 CC immobilised first polypeptide with the second polypeptide; and assaying
 CC the modification. The immobilised polypeptide and/or the binding partner
 CC polypeptide is/are associated with a label such as a fluorescent label or
 CC a radioactive label. The detectable signal is generated by an interaction
 CC between the labels comprising energy transfer. The association is
 CC measured by monitoring the molecular mass of the hybrid species
 CC comprising the second polypeptide associated with the first polypeptide,
 CC by surface plasma resonance, by scintillation proximity assay, or by
 CC using an antibody. Assaying the modification comprises assaying
 CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
 CC geranylactin, ubiquitination, prenylation, serratinsation, adenosine
 CC diophosphate (ADP)-ribosylation, or the reversal of any of these
 CC modifications. The method is used for analysing a sample, i.e. detecting
 CC or monitoring the activity of a modulator of a polypeptide modifying
 CC agent. The invention enables only one polypeptide to be susceptible to a
 CC single modification, allowing only one modification event per complex,
 CC thus resulting in increased output of the assay, as effectively every
 CC modification event which takes place will have an effect on the readout.
 CC This is the amino acid sequence of a template peptide containing
 CC modification sites affected by the binding of a partner peptide AAM00653.
 CC The invention provides a method for monitoring the modification state of
 CC the template peptide on binding of it's partner.
 CC
 XX
 SQ Sequence 43 AA;
 QY
 Db 3 OKAAELNOKSKLEEQ 18
 14 QEIAOLEQENRLEOE 29

FT /label= 4heptad
 FT Peptide 48..94
 FT /label= F-zipper
 XX
 XX MO9705249-A2.
 XX
 XX 13-FEB-1997.
 PD
 XX
 PF 31-JUL-1996; 96MO-US012590.
 XX
 XX 31-JUL-1995; 95US-0001654P.
 XX 29-MAY-1996; 96US-0018496P.
 PR 31-JUL-1996; 96US-00690011.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Vinson CR, Krylov D;
 XX
 XX WPI; 1997-145687/13.
 DR N-PSDB; AAT64355.
 DR
 XX New nucleic acid binding proteins - having an acidic amino acid sequence
 PT extension at the amino-terminus, to increase ability to regulate gene
 PT transcription, useful e.g. in cancer therapeutics.
 XX
 PS Example 1; Page 96; 144pp; English.
 XX
 CC This polypeptide sequence comprises an N-terminal leader sequence, 3
 CC glycine residues which, when deleted, produce a more effective dominant
 CC negative, a 4heptad sequence (see also AAM00961) and an F zipper (see
 CC also AAM00957). It was designed for constructs that include 4heptad-F
 CC zipper protein. Claimed nucleic acid binding proteins (NABPs), including
 CC F zipper proteins, have acidic peptide extensions are capable of
 CC regulating the function of a target nucleic acid or gene to which they
 CC are bound, and act as potent dominant-negative regulators of gene
 CC transcription, cell growth and cell proliferation. They can be used in
 CC cancer therapeutics, to treat diseases caused by eukaryotic
 CC microorganisms or by viruses, and as tools for drug development, rational
 CC drug design, and drug and gene therapies. They have an extended protein
 CC interaction surface or multimerisation or dimerisation interface that
 CC increases the stability of the complexes formed
 CC
 XX
 SQ Sequence 94 AA;
 QY
 Db 3 OKAAELNOKSKLEEQ 18
 34 KEAEELQENAELEOE 49

RESULT 24
 AAM00956
 ID AAM00956 standard; protein; 94 AA.
 XX
 AC AAM00956;
 XX
 DT 12-NOV-1997 (first entry)
 XX
 DE 4heptad-F zipper protein.
 XX
 KW DNA binding protein; RNA binding protein; amphipathic peptide;
 KW acidic extension peptide; gene control; gene regulation; transcription;
 KW dominant negative protein; cancer; drug therapy; drug design.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH 1..14
 FT Peptide /label= Leader_peptide
 FT 18..47
 FT Peptide

RESULT 25
 AAB08375
 ID AAB08375 standard; peptide; 43 AA.
 XX
 AC AAB08375;
 XX
 DT 20-DEC-2000 (first entry)
 XX
 DE Immobilised substrate polypeptide used to assay for S6 kinase activity.
 XX
 KW Binding partner; protein modification; post-translational modification;
 KW modulator; coiled-coil structure.
 XX
 OS Unidentified.
 XX
 XX WO200050902-A2.
 XX
 XX 31-AUG-2000.
 PD
 XX 25-FEB-2000; 2000MO-GB000669.
 XX

XX 25-FEB-1999; 99GB-00004398.
XX (FLUO-) FLUORESCIENCE LTD.
XX Colyer J, Craig RK, Maschio A, Mezna M;
XX WPI; 2000-572119/53.
XX
XX High throughput assay for monitoring modification of polypeptides and
XX PT modulation of the modifications.
XX
XX Example 5; Page 76; 128pp; English.
XX
XX The specification describes a method for analysing a sample. The method
XX comprises immobilising a polypeptide to a physical support, contacting
XX the immobilised polypeptide with a test sample which may contain an agent
XX capable of modifying the immobilised polypeptide, contacting the
XX immobilised polypeptide with a binding partner polypeptide, where
XX association of both polypeptide is dependent on the modification state of
XX the immobilised polypeptide, and measuring the association of the binding
XX partner polypeptide to the immobilised polypeptide. The polypeptides,
XX support and methods can be used to analyse a sample to determine if
XX modification of a polypeptide is taking place and to identify modulators
XX of the modification. This is useful for monitoring the post-translational
XX modification of proteins. The present sequence represents an immobilised
XX substrate polypeptide, which is used in the method of the invention, to
XX assay for S6 kinase activity
XX
XX Sequence 43 AA;
XX
XX Query Match 45.5%; Score 40; DB 3; Length 43;
XX Best Local Similarity 50.0%; Pred. No. 86;
XX Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
XX
XX 3 QKAAELNQSKKELEQQ 18
XX 14 QETARLRQESAQLEQQ 29
XX
XX
XX RESULT 26
XX ADA00682
XX ID ADA00682 standard; peptide; 43 AA.
XX
XX ADA00682;
XX
XX 06-NOV-2003 (first entry)
XX
XX Protein modification analysis method associated template peptide #25.
XX
XX protein modification analysis method; protein modifying agent;
XX surface plasma resonance; scintillation proximity assay; proteolysis;
XX phosphorylation; acylation; glycosylation; farnesylation; geranylation;
XX ubiquitination; prenylation; sentrimination;
XX adenosine diphosphate ribosylation; ADP ribosylation;
XX protein modification.
XX
XX Unidentified.
XX
XX US2003032054-A1.
XX
XX 13-FEB-2003.
XX
XX 26-FEB-1999; 99US-00259658.
XX
XX 26-FEB-1999; 99US-00259658.
XX
XX (WILL/) WILLIAMS K M.
XX
XX Colyer J, Craig RK, Maschio A, Mezna M;
XX WPI; 2003-605702/57.
XX

PT Analysis method, useful for detecting or monitoring the activity of a
PT modulator of a polypeptide modifying agent, comprises immobilizing a
PT polypeptide to a physical support, contacting with another polypeptide
PT and assaying.
XX
XX Example 5; Page 27; 32pp; English.
XX
XX The invention describes an analysis method, useful for detecting or
XX monitoring the activity of a modulator of a polypeptide modifying agent.
XX The method comprises: providing a polypeptide pair, where association of
XX the polypeptides is detectable and modification of at least one
XX polypeptide results in modulation of the association; contacting
XX immobilised first polypeptide with the second polypeptide; and assaying
XX the modification. The immobilised polypeptide and/or the binding partner
XX polypeptide is/are associated with a label such as a fluorescent label or
XX a radioactive label. The detectable signal is generated by an interaction
XX between the labels comprising energy transfer. The association is
XX measured by monitoring the molecular mass of the hybrid species
XX comprising the second polypeptide associated with the first polypeptide,
XX by surface plasma resonance, by scintillation proximity assay, or by
XX using an antibody. Assaying the modification comprises assaying
XX proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
XX geranylation, ubiquitination, prenylation, sentrimination, adenosine
XX diphosphate (ADP)-ribosylation, or the reversal of any of these
XX modifications. The method is used for analysing a sample, i.e. detecting
XX or monitoring the activity of a modulator of a polypeptide modifying
XX agent. The invention enables only one polypeptide to be susceptible to a
XX single modification, allowing only one modification event per complex,
XX thus resulting in increased output of the assay, as effectively every
XX modification event which takes place will have an effect on the readout.
XX This is the amino acid sequence of a template peptide containing
XX modification sites affected by the binding of a partner peptide ADA00653.
XX The invention provides a method for monitoring the modification state of
XX the template peptide on binding of it's partner.
XX
XX Sequence 43 AA;
XX
XX Query Match 45.5%; Score 40; DB 6; Length 43;
XX Best Local Similarity 50.0%; Pred. No. 86;
XX Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
XX
XX 3 QKAAELNQSKKELEQQ 18
XX 14 QETARLRQESAQLEQQ 29
XX
XX
XX RESULT 27
XX AAG01789
XX ID AAG01789 standard; protein; 82 AA.
XX
XX AAG01789;
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein, SEQ ID NO: 5870.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping.
XX
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-00200610.
XX
XX 26-FEB-1999; 99US-0122487P.
XX
XX (GERS) GENSERT.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX

DR WPI: 2000-500381/45.
DR N-PSDB: AAC01795.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS Claim 13; SEQ ID NO 5870; 71pp + Sequence Listing; English.
XX
CC The present sequence is a polypeptide encoded by one of a large number of
CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
CC prepared from total human RNAs or polyA+ RNAs derived from 30 different
CC tissues. EST sequences usually correspond mainly to the 3' untranslated
CC region (UTR) of the mRNA because they are often obtained from oligo-dT
CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
CC sequences derived from the 5' ends of mRNAs and even in those cases where
CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
CC are also used in diagnostic, forensic, gene therapy and chromosome
CC mapping procedures. They are used to obtain upstream regulatory sequences
CC and to design expression and secretion vectors
XX
SQ Sequence 82 AA;

Query Match 45.5%; Score 40; DB 3; Length 82;
Best Local Similarity 43.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

OY 3 OKAAELNOKSKELREQ 18
DB 60 KKQELNKKAEELDR 75

RESULT 28
AAG79886
ID AAG79886 standard; peptide; 46 AA.
XX
AC AAG79886;
XX
DT 28-APR-2003 (first entry)
XX
DE LOX-1 isoform 1 repeat #1.
XX
KW Mammalian; LOX-1; oxidised low density lipoprotein receptor; isoform;
KW atherosclerosis; heart failure; stroke; inflammation.
XX
OS Murine sp.
XX
PN WO2003004680-A2.
XX
PD 16-JAN-2003.
XX
PF 02-JUL-2002; 2002WO-US021075.
XX
PR 02-JUL-2001; 2001US-00898554.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Tail AR, Welch CL, Liang C;
XX
DR WPI; 2003-210370/20.
XX
PT New atherosclerosis susceptibility gene locus 1 and 2 (ATHSQ1 and ATHSQ2)
PT encoding a mammalian LOX-1 receptor protein, useful for treating an
PT abnormality such as atherosclerosis, heart failure, stroke, or
PT inflammation.
XX
PS Disclosure; Fig 13; 128pp; English.
XX
CC The sequences given in AAG79886-94 represent repeat regions derived from
CC mammalian LOX-1 (oxidised low density lipoprotein receptor) isoforms. LOX
CC -1 is a membrane bound receptor protein. Cells expressing compounds which
CC activate or inhibit the activity of LOX-1. The LOX-1 nucleic acids and
CC proteins, and identified compounds are useful for treating an abnormality
CC such as atherosclerosis, heart failure, stroke, or inflammation. LOX-1
CC isoforms 2, 5 and 6 have the same amino acid sequence
XX
SQ Sequence 46 AA;

Query Match 44.9%; Score 39.5; DB 6; Length 46;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

OY 2 TOKAAELNOKSKELREQ 17
DB 14 TOK---LNKSKQOE 26

RESULT 29
AAG79891
ID AAG79891 standard; peptide; 46 AA.
XX
AC AAG79891;
XX
DT 28-APR-2003 (first entry)
XX
DE LOX-1 isoform 4 repeat #1.
XX
KW Mammalian; LOX-1; oxidised low density lipoprotein receptor; isoform;
KW atherosclerosis; heart failure; stroke; inflammation.
XX
OS Murine sp.
XX
PN WO2003004680-A2.
XX
PD 16-JAN-2003.
XX
PF 02-JUL-2002; 2002WO-US021075.
XX
PR 02-JUL-2001; 2001US-00898554.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Tail AR, Welch CL, Liang C;
XX
DR WPI; 2003-210370/20.
XX
PT New atherosclerosis susceptibility gene locus 1 and 2 (ATHSQ1 and ATHSQ2)
PT encoding a mammalian LOX-1 receptor protein, useful for treating an
PT abnormality such as atherosclerosis, heart failure, stroke, or
PT inflammation.
XX
PS Disclosure; Fig 13; 128pp; English.
XX
CC The sequences given in AAG79886-94 represent repeat regions derived from
CC mammalian LOX-1 (oxidised low density lipoprotein receptor) isoforms. LOX
CC -1 is a membrane bound receptor protein. Cells expressing compounds which
CC activate or inhibit the activity of LOX-1. The LOX-1 nucleic acids and
CC proteins, and identified compounds are useful for treating an abnormality
CC such as atherosclerosis, heart failure, stroke, or inflammation. LOX-1
CC isoforms 2, 5 and 6 have the same amino acid sequence
XX
SQ Sequence 46 AA;

Query Match 44.9%; Score 39.5; DB 6; Length 46;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

OY 2 TOKAAELNOKSKELREQ 17
DB 14 TOK---LNKSKQOE 26

CC used in the method of the invention for identifying compounds which
CC activate or inhibit the activity of LOX-1. The LOX-1 nucleic acids and
CC proteins, and identified compounds are useful for treating an abnormality
CC such as atherosclerosis, heart failure, stroke, or inflammation. LOX-1
CC isoforms 2, 5 and 6 have the same amino acid sequence
XX
SQ Sequence 46 AA;

Query Match 44.9%; Score 39.5; DB 6; Length 46;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

OY 2 TOKAAELNOKSKELREQ 17
DB 14 TOK---LNKSKQOE 26

RESULT 29
AAG79891
ID AAG79891 standard; peptide; 46 AA.
XX
AC AAG79891;
XX
DT 28-APR-2003 (first entry)
XX
DE LOX-1 isoform 4 repeat #1.
XX
KW Mammalian; LOX-1; oxidised low density lipoprotein receptor; isoform;
KW atherosclerosis; heart failure; stroke; inflammation.
XX
OS Murine sp.
XX
PN WO2003004680-A2.
XX
PD 16-JAN-2003.
XX
PF 02-JUL-2002; 2002WO-US021075.
XX
PR 02-JUL-2001; 2001US-00898554.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Tail AR, Welch CL, Liang C;
XX
DR WPI; 2003-210370/20.
XX
PT New atherosclerosis susceptibility gene locus 1 and 2 (ATHSQ1 and ATHSQ2)
PT encoding a mammalian LOX-1 receptor protein, useful for treating an
PT abnormality such as atherosclerosis, heart failure, stroke, or
PT inflammation.
XX
PS Disclosure; Fig 13; 128pp; English.
XX
CC The sequences given in AAG79886-94 represent repeat regions derived from
CC mammalian LOX-1 (oxidised low density lipoprotein receptor) isoforms. LOX
CC -1 is a membrane bound receptor protein. Cells expressing compounds which
CC activate or inhibit the activity of LOX-1. The LOX-1 nucleic acids and
CC proteins, and identified compounds are useful for treating an abnormality
CC such as atherosclerosis, heart failure, stroke, or inflammation. LOX-1
CC isoforms 2, 5 and 6 have the same amino acid sequence
XX
SQ Sequence 46 AA;

Query Match 44.9%; Score 39.5; DB 6; Length 46;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

OY 2 TOKAAELNOKSKELREQ 17
DB 14 TOK---LNKSKQOE 26

CC -1 is a membrane bound receptor protein. Cells expressing LOX-1 may be

```

RESULT 30
AAG79889
ID AAG79889 standard; peptide: 46 AA.
XX
XX
AC AAG79889;
XX
DT 28-APR-2003 (first entry)
XX
DE LOX-1 isoform 3 repeat #1.
XX
XX Mammalian; LOX-1; oxidised low density lipoprotein receptor; isoform;
KM atherosclerosis; heart failure; stroke; inflammation.
XX
OS Murine sp.
XX
PN W02003004680-A2.
XX
PD 16-JAN-2003.
XX
PF 02-JUL-2002; 2002WO-US021075.
XX
PR 02-JUL-2001; 2001US-00898554.
XX
PA (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI Tall AR, Welch CL, Liang C;
XX
DR WPI; 2003-210370/20.
XX
PT New atherosclerosis susceptibility gene locus 1 and 2 (ATHSQ1 and ATHSQ2)
PT encoding a mammalian LOX-1 receptor protein, useful for treating an
PT abnormality such as atherosclerosis, heart failure, stroke, or
PT inflammation.
XX
PS Disclosure; Fig 13; 128pp; English.
XX
XX The sequences given in AAG79886-94 represent repeat regions derived from
CC mammalian LOX-1 (oxidised low density lipoprotein receptor) isoforms. LOX
CC -1 is a membrane bound receptor protein. Cells expressing LOX-1 may be
CC used in the method of the invention for identifying compounds which
CC activate or inhibit the activity of LOX-1. The LOX-1 nucleic acids and
CC proteins, and identified compounds are useful for treating an abnormality
CC such as atherosclerosis, heart failure, stroke, or inflammation. LOX-1
CC isoforms 2, 5 and 6 have the same amino acid sequence
XX
SQ Sequence 46 AA;
XX
Query Match 44.9%; Score 39.5; DB 6; Length 46;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 1; Indels 3; Gaps 1;
QY 2 TOKAAELNOKSKLEEQ 17
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |
Db 14 TOK--LNEKSKQEE 26
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |

RESULT 31
AAB37026
ID AAB37026 standard; peptide: 27 AA.
XX
XX
AC AAB37026;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #26.
XX
XX Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KM cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KM apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KM colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KM melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS; stroke;
KM myocardial infarction.
XX

```

```

OS Homo sapiens.
XX
XX W0200059526-A1.
XX
XX
PD 12-OCT-2000.
XX
XX
PF 06-APR-2000; 2000WO-US009352.
XX
XX
PR 07-APR-1999; 99US-0128202P.
XX
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
XX Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
XX WPI; 2000-679325/66.
XX
XX
DR
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B cell
PT lymphoma/leukemia 2 (Bcl-2) function, especially useful for treating
PT neurodegenerative disorders, stroke, or cancer.
XX
XX
PS Claim 18; Page 18; 74pp; English.
XX
XX The invention relates to a peptide conjugate having the formula: (R-X)n-
CC peptide where n = 1-10; X = C=O, when the R-X group is attached to the N-
CC terminus of the peptide, or a side chain of the peptide where the
CC functional group of the side chain is NH2 or OH; or X = O or NH, when the
CC R-X group is attached to the C-terminus of the peptide, or a side chain
CC of the peptide, where the side chain functional group is COOH or CONH2;
CC and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one or two
CC double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric, non-
CC small lung, renal or thyroid cancers, neuroblastoma, melanoma, or acute
CC or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction
XX
XX
SQ Sequence 27 AA;
XX
Query Match 44.3%; Score 39; DB 3; Length 27;
Best Local Similarity 44.4%; Pred. No. 74;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 1 PTOKAELNOKSKLEEQ 18
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |
Db 5 PSATPAELRRRAAELERR 22
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |

RESULT 32
AAB08379
ID AAB08379 standard; peptide: 31 AA.
XX
XX
AC AAB08379;
XX
DT 20-DEC-2000 (first entry)
XX
XX
DE Peptide used to assay for peptidase activity.
XX
XX Binding partner; protein modification; post-translational modification;
KM modulator; coiled-coil structure.
XX
XX
XX Unidentified.
XX

```



```

DR WP1; 2000-572119/53.
XX High throughput assay for monitoring modification of polypeptides and
PT modulation of the modifications.
XX
PS Example 21; Page 100; 128pp; English.
XX
CC The specification describes a method for analysing a sample. The method
CC comprises immobilising a polypeptide to a physical support, contacting
CC the immobilised polypeptide with a test sample which may contain an agent
CC capable of modifying the immobilised polypeptide, contacting the
CC immobilised polypeptide with a binding partner polypeptide, where
CC association of both polypeptide is dependent on the modification state of
CC the immobilised polypeptide, and measuring the association of the binding
CC partner polypeptide to the immobilised polypeptide. The polypeptides,
CC support and methods can be used to analyse a sample to determine if
CC modification of a polypeptide is taking place and to identify modulators
CC of the modification. This is useful for monitoring the post-translational
CC modification of proteins. The present sequence represents a peptide,
CC which may be used in the method of the invention, to assay for Src kinase
CC activity
CC
SQ Sequence 31 AA;
OY Query Match 44.3%; Score 39; DB 3; Length 31;
Best Local Similarity 43.8%; Pred. No. 86;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
DB 3 OKAAELNOKSKETLEEQ 18
| : | : | : | : | : | : | : | : | :
9 QEIRKQLEDENNAQLRQE 24
RESULT 34
ABU09302
ID ABU09302 standard; peptide; 31 AA.
AC ABU09302;
DT 25-JUN-2003 (first entry)
XX
DE Peptide 4HA used in method for monitoring enzyme activity.
XX
KW Enzyme activity; binding partner polypeptide; enzyme modulator; kinase;
KW phosphatase; O-glucanC transferase; carboxydrate transferase;
KW UDP-N-acetylglucosamine-dolichyl-phosphate-N-acetylglycosamine;
KW phosphotransferase; glycyIpeptide-N-tetradecanoyl transferase;
KW ubiquitin activating enzyme E1; ubiquitin conjugating enzyme E2;
KW ubiqutin conjugating enzyme Ubc9; ubiquitin protein Ligase E3;
KW poly(ADP-ribose) polymerase; fatty acyl transferase;
KW NAD:arginine ADP ribosyltransferase; protein modifying enzyme.
OS Synthetic.
XX
PN US2002197606-A1.
PD 26-DEC-2002.
PF 25-JAN-2001; 2001US-00770102.
PR 31-JAN-2000; 2000US-0179283P.
PA (CRAI/) CRAIG R.
PI Craig R;
WP1; 2003-361860/34.
XX
DR Monitoring activity of enzymes e.g. kinase, as a function of
PT association/dissociation of binding partner polypeptides dependent upon
PT on addition or removal of group to one or more binding partner
PT polypeptides by enzyme.

```

PS Example 1; Page 27; 44pp; English..

The present invention relates to compositions and methods for monitoring the activity of one or more enzymes. The method comprising detecting binding or dissociation of one or more binding partner polypeptides with the corresponding one or more tagged binding partner polypeptides, in the presence of one or more enzymes that add or remove a group to or from one or more binding partner polypeptide. The method is useful for screening for a candidate modulator of enzymatic activity of the following enzymes: kinases, phosphatases, UDP-N-acetylglucosamine-6-olcyl-1-phosphate-N-acetylglucosamine phosphotransferase, O-6-olcNAc transferase, glycyloleptidase, N-tetradecanoyl transferase, carboxylate transferase, ubiquitin activating enzyme E1, ubiquitin conjugating enzyme E2, poly(ADP-ribose) conjugating enzyme Ubc9, ubiquitin protein ligase E3, poly(ADP-ribose) polymerase, fatty acyl transferase, and NAD-arginine ADP ribosyltransferase. The method of the invention is rapid, and quantitatively measures the protein modifying enzymes, by measuring the extent of binding of isolated preparations of binding domains in response to incubation with a modifying enzyme. Using combinations of different modification dependent binding partner polypeptides, the activities of several enzymes can be measured simultaneously. The present invention represents a peptide used in the examples of the present invention

SQ Sequence 31 AA;

Query Match	44.3%	Score 39;	DB 6;	Length 31;
Best Local Similarity	43.8%	Pred. No. 86;		
Matches 7; Conservative	6;	Mismatches	0;	Gaps 0;

QY 3 QKAÆLNÖKSKELEÖÖ 18

Db 9 QEI AQL E QNA QLE Q 24

RESULT 35
ADA00689

AC ADA00689;

DT 06-NOV-2003 (first entry)

DE Protein modification analysis method associated template peptide #29.

protein modification analysis method; protein modifying agent;
surface plasma resonance; scintillation proximity assay; proteolysis;
phosphorylation; acylation; glycosylation; farnesylation; geranyl-
ubiquitination; prenylation; sentrinisation;
adenosine diphosphate ribosylation; ADP ribosylation;
protein modification, yeast; human.

Homo sapiens.

PN US2003032054-A1.

PD 13-FEB-2003

PF 26-FEB-1999; 99US-00259658.

PR 26-FEB-1999; 99US-00259658.

PA (WILL/) WILLIAMS K M.

PI Colyer J, Craig RK, Maschio A, Mezna M;

DR WPI; 2003-605702/57.

Analysis method, useful for detecting or monitoring the activity of a modulator of a polypeptide modifying agent, comprises immobilising a polypeptide to a physical support, contacting with another polypeptide and assaying.

PS Example 8; Page 28; 32pp; English.

XX The invention describes an analysis method, useful for detecting or monitoring the activity of a modulator of a polypeptide modifying agent. The method comprises: providing a polypeptide pair, where association of the polypeptides is detectable and modification of at least one polypeptide results in modulation of the association; contacting immobilised first polypeptide with the second polypeptide; and assaying the modification. The immobilised polypeptide and/or the binding partner polypeptide is/are associated with a label such as a fluorescent label or a radioactive label. The detectable signal is generated by an interaction between the labels comprising energy transfer. The association is measured by monitoring the molecular mass of the hybrid species comprising the second polypeptide associated with the first polypeptide, by surface plasma resonance, by scintillation proximity assay, or by using an antibody. Assaying the modification comprises assaying proteolysis, phosphorylation, acylation, glycosylation, farnesylation, geranylolation, ubiquitination, prenylation, serinylation, adenosine diphosphate (ADP)-ribosylation, or the reversal of any of these modifications. The method is used for analysing a sample, i.e. detecting or monitoring the activity of a modulator of a polypeptide modifying agent. The invention enables only one polypeptide to be susceptible to a single modification, allowing only one modification event per complex, thus resulting in increased output of the assay, as effectively every modification event which takes place will have an effect on the readout. This is the amino acid sequence of a template peptide containing modification sites affected by the binding of a partner peptide ADO0653. The invention provides a method for monitoring the modification state of the template peptide on binding of it's partner.

SQ Sequence 33 AA;

Query Match	44.3%	Score 39;	DB 6;	Length 33;
Best Local Similarity	43.8%;	Pred. No. 92;		
Matches 7; Conservative	6;	Mismatches	3;	Indels 0;
				Gaps 0;

QY 3 QKAEINQSKELJQQ 18

Db 14 QEIΛQLEQENAQLQE 29

RESULT 36
ADM97089

XX

XX: 1991

XX

[illegible]

XX	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54	57	60	63	66	69	72	75	78	81	84	87	90	93	96	99	102	105	108	111	114	117	120	123	126	129	132	135	138	141	144	147	150	153	156	159	162	165	168	171	174	177	180	183	186	189	192	195	198	201	204	207	210	213	216	219	222	225	228	231	234	237	240	243	246	249	252	255	258	261	264	267	270	273	276	279	282	285	288	291	294	297	300	303	306	309	312	315	318	321	324	327	330	333	336	339	342	345	348	351	354	357	360	363	366	369	372	375	378	381	384	387	390	393	396	399	402	405	408	411	414	417	420	423	426	429	432	435	438	441	444	447	450	453	456	459	462	465	468	471	474	477	480	483	486	489	492	495	498	501	504	507	510	513	516	519	522	525	528	531	534	537	540	543	546	549	552	555	558	561	564	567	570	573	576	579	582	585	588	591	594	597	600	603	606	609	612	615	618	621	624	627	630	633	636	639	642	645	648	651	654	657	660	663	666	669	672	675	678	681	684	687	690	693	696	699	702	705	708	711	714	717	720	723	726	729	732	735	738	741	744	747	750	753	756	759	762	765	768	771	774	777	780	783	786	789	792	795	798	801	804	807	810	813	816	819	822	825	828	831	834	837	840	843	846	849	852	855	858	861	864	867	870	873	876	879	882	885	888	891	894	897	900	903	906	909	912	915	918	921	924	927	930	933	936	939	942	945	948	951	954	957	960	963	966	969	972	975	978	981	984	987	990	993	996	999
XX	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54	57	60	63	66	69	72	75	78	81	84	87	90	93	96	99	102	105	108	111	114	117	120	123	126	129	132	135	138	141	144	147	150	153	156	159	162	165	168	171	174	177	180	183	186	189	192	195	198	201	204	207	210	213	216	219	222	225	228	231	234	237	240	243	246	249	252	255	258	261	264	267	270	273	276	279	282	285	288	291	294	297	300	303	306	309	312	315	318	321	324	327	330	333	336	339	342	345	348	351	354	357	360	363	366	369	372	375	378</																																																																																																																																																																																																															

XX 8

[illegible]

XXV

XX

XX

XX

XX

XX Determining clostridial toxin activity, comprises contacting cell with
PT sample comprising substrate with donor fluorophore, acceptor and
PT recognition sequence, exciting fluorophore, and determining resonance
PT energy transfer of contacted cell.

PS Disclosure; SEQ ID NO 95; 188pp; English.

XX The invention relates to a method of determining (M1) clostridial toxin

CC activity, comprising contacting cell with sample comprising clostridial

CC toxin substrate with donor fluorophore (F), acceptor (A) with an

CC absorbance spectrum with overlapping emission spectrum of fluorophore

CC clostridial toxin recognition sequence with cleavage site intervening

CC between (F) and (A), where under appropriate conditions resonance

CC transfer is exhibited between fluorophore and acceptor, exciting the

CC donor fluorophore, and determining resonance energy transfer of the

CC contacted cell relative to a control cell, where a difference in

CC resonance energy transfer of the contacted cell as compared to the

CC control cell is indicative of clostridial toxin activity. (M1) is useful

CC for determining clostridial toxin activity, where the sample is a crude

CC cell lysate, isolated clostridial toxin, formulated clostridial toxin

CC product, BOTOX or food. (M1) is an automated high-throughput assay. (M1)

CC reduces the need for animal toxicity studies and serves to analyze

CC multiple toxin functions such as binding a cellular uptake of the toxin,

CC translocation into the cell cytosol and protease activity. In the method

CC M1, the clostridial toxin substrate is a botulinum toxin substrate

CC selected from a BONT/A, BONT/B, BONT/C1, BONT/D, BONT/E, BONT/F, BONT/G

CC or TONT substrate comprising a BONT/A, BONT/B, BONT/C1, BONT/D, BONT/E,

CC BONT/F, BONT/G or TONT recognition sequence. This sequence corresponds to

CC a recognition sequence for a botulinum toxin used as the clostridial

CC toxin substrate.

XX Sequence 40 AA:

SO

Query Match 44.3%; Score 39; DB 8; Length 40;

Best Local Similarity 46.7%; Pred. No. 1.1e+02;

Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 OKAAELNOKSKKEEQ 17

DB 8 QKSELGERADQLEQ 22

RESULT 37

AAB08344

ID AAB08344 standard; peptide; 43 AA.

XX AAB08344;

AC

DT 20-DEC-2000 (first entry)

XX

DE Amino acid sequence of a coiled-coil structure template.

XX

KW Binding partner; protein modification; post-translational modification;

KM modulator; coiled-coil structure.

XX

OS Unidentified.

XX

PN WO200050902-A2.

PD 31-AUG-2000.

XX

PF 25-FEB-2000; 2000WO-GB000669.

XX

PR 25-FEB-1999; 99GB-00004398.

XX

PA (FLUO-) FLUORESCENCE LTD.

XX

PI Colyer J, Craig RK, Maschio A, Mezna M;

XX

DR WPI; 2000-572119/53.

XX

PT High throughput assay for monitoring modification of polypeptides and

PT modulation of the modifications.

XX

PS Disclosure; Page 21; 128pp; English.

XX

CC The specification describes a method for analysing a sample. The method

CC comprises immobilising a polypeptide to a physical support, contacting

CC the immobilised polypeptide with a test sample which may contain an agent

CC capable of modifying the immobilised polypeptide, contacting the

CC immobilised polypeptide with a binding partner polypeptide, where

CC association of both polypeptide is dependent on the modification state of

CC the immobilised polypeptide, and measuring the association of the binding

CC partner polypeptide to the immobilised polypeptide. The polypeptides,

CC support and methods can be used to analyse a sample to determine if

CC modification of a polypeptide is taking place and to identify modulators

CC of the modification. This is useful for monitoring the post-translational

CC modification of proteins. AAB08344-45 represent binding partners which

CC have coiled-coil structures, and may be used in the method of the

CC invention

XX

SO Sequence 43 AA;

QY 3 OKAAELNOKSKKEEQ 18

DB 14 QETALQLEQENALQLEQ 29

RESULT 38

ADA00652

ID ADA00652 standard; peptide; 43 AA.

XX ADA00652;

AC

DT 06-NOV-2003 (first entry)

XX

DE Protein modification analysis method associated template peptide #1.

XX

KW protein modification analysis method; protein modifying agent;

KM surface plasma resonance; scintillation proximity assay; proteolysis;

KW phosphorylation; acylation; glycosylation; farnesylation; geranylation;

KW ubiquitination; prenylation; sentrinisation;

KM adenosine diphosphate ribosylation; ADP ribosylation;

XX

OS Unidentified.

XX

PN US2003032054-A1.

PD 13-FEB-2003.

XX

PF 26-FEB-1999; 99US-00259658.

XX

PR 26-FEB-1999; 99US-00259658.

XX

PA (WILL/) WILLIAMS K M.

XX

PI Colyer J, Craig RK, Maschio A, Mezna M;

XX

DR WPI; 2003-605702/57.

XX

PT Analysis method, useful for detecting or monitoring the activity of a

PT modulator of a polypeptide modifying agent, comprises immobilizing a

PT polypeptide to a physical support, contacting with another polypeptide

PT and assaying.

XX

PS Disclosure; Page 8; 32pp; English.

XX

CC The invention describes an analysis method, useful for detecting or

CC monitoring the activity of a modulator of a polypeptide modifying agent.

CC The method comprises: providing a polypeptide pair, where association of

CC the polypeptides is detectable and modification of at least one

CC polypeptide results in modulation of the association; contacting

CC immobilised first polypeptide with the second polypeptide; and assaying

CC the modification. The immobilised polypeptide and/or the binding partner

CC polypeptide is/are associated with a label such as a fluorescent label or

CC a radioactive label. The detectable signal is generated by an interaction

CC between the labels comprising energy transfer. The association is
 CC measured by monitoring the molecular mass of the hybrid species
 CC comprising the second polypeptide associated with the first polypeptide,
 CC by surface plasma resonance, by scintillation proximity assay, or by
 CC using an antibody. Assaying the modification comprises assaying
 CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
 CC geranylation, ubiquitination, prenylation, serinisation, adenosine
 CC diphosphate (ADP)-ribosylation, or the reversal of any of these
 CC modifications. The method is used for analysing a sample, i.e. detecting
 CC or monitoring the activity of a modulator of a polypeptide modifying
 CC agent. The invention enables only one polypeptide to be susceptible to a
 CC single modification, allowing only one modification event per complex,
 CC thus resulting in increased output of the assay, as effectively every
 CC modification event which takes place will have an effect on the readout.
 CC This is the amino acid sequence of a template peptide containing
 CC modification sites affected by the binding of a partner peptide AAM00653.
 CC The invention provides a method for monitoring the modification state of
 CC the template peptide on binding of it's partner.

XX Sequence 43 AA;

SO Query Match 44.3%; Score 39; DB 6; Length 43;
 Best Local Similarity 43.8%; Pred. No. 1.2e+02;
 Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Oy 3 OKAAELNOKSEKEEQ 18
 Db 14 QEIAPLEQENAELEAE 29

RESULT 39

AAM00949
 ID AAM00949 standard; protein; 97 AA.

XX AC AAM00949;

XX DT 12-NOV-1997 (first entry)

XX DE CMV500-4heptadFos leucine zipper protein.

XX KW DNA binding protein; RNA binding protein; amphipathic peptide;
 KW acidic extension peptide; gene control; gene regulation; transcription;
 KW dominant negative protein; c-Fos; cancer; drug therapy; drug design.

XX OS Homo; sapiens.
 OS unidentified bacteriophage; phi-10.
 OS Synthetic.
 OS Chimeric.

XX FH Key Location/Qualifiers
 FT Peptide 1
 FT Protein /label= FLAG
 FT Protein 11..22
 FT Protein /label= Phi-10
 FT Protein 24..53
 FT Protein /label= 4Heptad_extension
 FT Protein 55..97
 FT Protein /label= C-Fos

XX MO9705249-A2.

XX PD 13-FEB-1997.

XX PF 31-JUL-1996; 96WO-US012590.

XX PR 31-JUL-1995; 95US-0001654P.

XX PR 29-MAY-1996; 96US-0018496P.

XX PR 31-JUL-1996; 96US-00690011.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Vinson CR, Krylov D;

XX

DR WPI; 1997-145687/13.

DR N-PSDB; AAT84348.

XX New nucleic acid binding proteins - having an acidic amino acid sequence
 PT extension at the amino-terminus, to increase ability to regulate gene
 PT transcription, useful e.g. in cancer therapeutics.

XX Claim 16; Page 88; 144pp; English.

XX This protein sequence comprises CMV500-4heptadFos leucine zipper, a
 CC modified Fos nucleic acid binding protein (NABP) that includes an
 CC appended 4-heptad acidic extension. Claimed NABPs such as Fos that have
 CC acidic peptide extensions are capable of regulating the function of a
 CC target nucleic acid or gene to which they are bound, and act as potent
 CC dominant-negative regulators of gene transcription, cell growth and cell
 CC proliferation. They can be used in cancer therapeutics, to treat diseases
 CC caused by eukaryotic microorganisms or by viruses, and as tools for drug
 CC development, rational drug design, and drug and gene therapies. They have
 CC an extended protein interaction surface or multimerisation or
 CC dimerisation interface that increases the stability of complexes formed

XX Sequence 97 AA;

SO Query Match 44.3%; Score 39; DB 2; Length 97;
 Best Local Similarity 50.0%; Pred. No. 3e+02;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy 3 OKAAELNOKSEKEEQ 18
 Db 42 KKAEELEQENAELEAE 57

RESULT 40

ADK17053
 ID ADK17053 standard; protein; 83 AA.

XX AC ADK17053;

XX DT 06-MAY-2004 (first entry)

XX DE Nanoarchaeum equitans cancer-associated (CA) protein #502.

XX KW cancer-associated gene; CA gene; cancer; carcinoma; lymphoma; leukaemia.

XX OS Nanoarchaeum equitans.

XX FN WO2003093434-A2.

XX PD 13-NOV-2003.

XX PF 01-MAY-2003; 2003WO-US013699.

XX PR 01-MAY-2002; 2002US-0377447P.

XX PA (DIVE-) DIVERSA CORP.

XX PI Stettler KO, Waters E, Kretz K, Podar M, Richardson T;

XX PI Noordewier M;

XX WPI; 2004-053041/05.

XX DR N-PSDB; ADK17052.

XX New recombinant cancer-associated genes, such as KCNJ9, useful for
 PT diagnosing or treating carcinoma, e.g. breast, colon, rectal, pancreatic,
 PT cervical, or skin cancers, lymphomas, or leukemia.

XX Claim 64; SEQ ID NO 1005; 251pp; English.

XX The invention comprises then amino acid and coding sequences of cancer-
 CC associated (CA) genes isolated from Nanoarchaeum equitans. The invention
 CC also comprises the Nanoarchaeum equitans genome. The DNA and protein
 CC sequences of the invention are useful for diagnosing and treating cancer
 CC (e.g. carcinoma, lymphoma, or leukemia). The present amino acid sequence

DE Peptide used to assay for geranylgeranyl transferase activity.

XX Binding partner; protein modification; post-translational modification;
KM modulator; coiled-coil structure.
XX Unidentified.
XX WO2000050902-A2.
XX 31-AUG-2000.
XX 25-FEB-2000; 2000WO-GB000669.
XX 25-FEB-1999; 99GB-00004398.
XX (FLUO-) FLUORESCENCE LTD.
XX Colyer J, Craig RK, Maschio A, Mezna M;
XX WPI; 2000-572119/53.
XX High throughput assay for monitoring modification of polypeptides and
PT modulation of the modifications.
XX Example 12; Page 84; 128pp; English.
XX The specification describes a method for analysing a sample. The method
CC comprises immobilising a polypeptide to a physical support, contacting
CC the immobilised polypeptide with a test sample which may contain an agent
CC capable of modifying the immobilised polypeptide, contacting the
CC immobilised polypeptide with a binding partner polypeptide, where
CC association of both polypeptide is dependent on the modification state of
CC the immobilised polypeptide, and measuring the association of the binding
CC partner polypeptide to the immobilised polypeptide. The polypeptides,
CC support and methods can be used to analyse a sample to determine if
CC modification of a polypeptide is taking place and to identify modulators
CC of the modification. This is useful for monitoring the post-translational
CC modification of proteins. AAB08387-88 represent binding partners which
CC are used, in the method of the invention, to assay for geranylgeranyl
CC transferase activity
XX
SQ Sequence 31 AA;
Query Match 43.2%; Score 38; DB 3; Length 31;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 3 QKAAELNOKSKELE 16
Db 14 QEIAQLQENKQKLE 27
RESULT 44
ADA00698
ID ADA00698 standard; peptide; 31 AA.
XX
AC ADA00698;
XX
XX 06-NOV-2003 (first entry)
XX Protein modification analysis method associated template peptide #34.
XX
XX protein modification analysis method; protein modifying agent;
KM surface plasma resonance; scintillation proximity assay; proteolysis;
KM phosphorylation; acylation; glycosylation; farnesylation; geranylation;
KM ubiquitination; prenylation; sentrinisation;
KM adenosine diphosphate ribosylation; ADP ribosylation;
KM protein modification; yeast.
XX
XX Synthetic.
XX US2003032054-A1.
PN
XX
PD 13-FEB-2003.

XX
XX 26-FEB-1999; 99US-00259658.
PF
XX 26-FEB-1999; 99US-00259658.
PR
XX (WILL) WILLIAMS R M.
XX
XX Colyer J, Craig RK, Maschio A, Mezna M;
XX WPI; 2003-605702/57.
XX
XX Analysis method, useful for detecting or monitoring the activity of a
PT modulator of a polypeptide modifying agent, comprises immobilizing a
PT polypeptide to a physical support, contacting with another polypeptide
PT and assaying.
XX
XX Example 12; Page 30; 32pp; English.
XX
XX The invention describes an analysis method, useful for detecting or
CC monitoring the activity of a modulator of a polypeptide modifying agent.
CC The method comprises: providing a polypeptide pair where association of
CC the polypeptides is detectable and modification of at least one
CC polypeptide results in modulation of the association; contacting
CC immobilised first polypeptide with the second polypeptide; and assaying
CC the modification. The immobilised polypeptide and/or the binding partner
CC polypeptide is/are associated with a label such as a fluorescent label or
CC a radioactive label. The detectable signal is generated by an interaction
CC between the labels comprising energy transfer. The association is
CC measured by monitoring the molecular mass of the hybrid species
CC comprising the second polypeptide associated with the first polypeptide,
CC by surface plasma resonance, by scintillation proximity assay, or by
CC using an antibody. Assaying the modification comprises assaying
CC proteolysis, phosphorylation, acylation, glycosylation, farnesylation,
CC geranylation, ubiquitination, prenylation, sentrinisation, adenosine
CC diphosphate (ADP)-ribosylation, or the reversal of any of these
CC modifications. The method is used for analysing a sample, i.e. detecting
CC or monitoring the activity of a modulator of a polypeptide modifying
CC agent. The invention enables only one polypeptide to be susceptible to a
CC single modification, allowing only one modification event per complex,
CC thus resulting in increased output of the assay, as effectively every
CC modification event which takes place will have an effect on the readout.
CC This is the amino acid sequence of a template peptide containing
CC modification sites affected by the binding of a partner peptide ADA00653.
CC The invention provides a method for monitoring the modification state of
CC the template peptide on binding of it's partner.
XX
SQ Sequence 31 AA;
Query Match 43.2%; Score 38; DB 6; Length 31;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 3 QKAAELNOKSKELE 16
Db 14 QEIAQLQENKQKLE 27
RESULT 45
AAW00961
ID AAW00961 standard; peptide; 33 AA.
XX
XX AAW00961;
AC
XX
DT 12-NOV-1997 (first entry)
XX
XX Heptad peptide 4heptad-F.
DE
XX DNA binding protein; RNA binding protein; amphipathic peptide;
KM acidic extension peptide; gene control; gene regulation; transcription;
KM dominant negative protein; cancer; drug therapy; drug design; F zipper;
KM heptad.
XX
XX Synthetic.

```

XX      Key      Location/Qualifiers
FH      Misc-difference 32
FT      /note="first L position of the F leucine zipper"
XX      WO9705249-A2.
XX      PD      13-FEB-1997.
XX      PP      31-JUL-1996; 96WO-US012590.
XX      PR      31-JUL-1995; 95US-0001654P.
XX      PR      29-MAY-1996; 96US-0018496P.
XX      PR      31-JUL-1996; 96US-00690011.
XX      PA      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX      PI      Vineson CR, Krylov D;
XX      DR      WPI; 1997-145687/13.
XX      PT      New nucleic acid binding proteins - having an acidic amino acid sequence
XX      PT      extension at the amino-terminus, to increase ability to regulate gene
XX      PT      transcription, useful e.g. in cancer therapeutics.
XX      PS      Claim 16; Page 98; 144pp; English.
XX      CC      This peptide sequence comprises a 4-heptad extension peptide that can be
XX      CC      appended to nucleic acid binding proteins (NABPs) in order to extend the
XX      CC      protein interaction surface or multimerisation or dimerisation interface
XX      CC      and hence to increase the stability of formed complexes. Claimed NABPs,
XX      CC      e.g. F zipper and bHLH proteins, having acidic extensions can regulate
XX      CC      the function of a target nucleic acid or gene to which they are bound,
XX      CC      and act as potent dominant-negative regulators of gene transcription,
XX      CC      cell growth and cell proliferation. They can be used in cancer
XX      CC      therapeutics, to treat diseases caused by eukaryotic microorganisms or by
XX      CC      viruses, and as tools for drug development, rational drug design, and
XX      CC      drug and gene therapies
XX      SQ      Sequence 33 AA;
XX      Query Match      43.2%; Score 38; DB 2; Length 33;
XX      Best Local Similarity 43.8%; Pred. No. 1.3e+02;
XX      Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
QY      3 QKAAELNOKSKELRQ 18
      6 QRAEELARENELEKE 21

```

Search completed: August 29, 2005, 17:01:42
 Job time : 167 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2004, 16:50:52 ; Search time 38 Seconds

(without alignments)
45.576 Million cell updates/sec

Title: US-10-031-289-1331

Perfect score: 88

Sequence: 1 PTKAAELNKSKELEQ 18

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Listing first 100 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	1975	2	B81192 hemagglutinin/hemo
2	88	100.0	1995	2	B81044 hemagglutinin/hemo
3	88	100.0	2015	2	B81989 hypothetical prote
4	49	55.7	2469	2	H36812 hypothetical prote
5	47	53.4	566	2	S19063 hypothetical prote
6	46	52.3	333	2	AE2084 ABC transporter, p
7	46	52.3	810	2	S67696 probable membrane
8	45	51.1	172	2	C72327 hypothetical prote
9	45	51.1	440	2	B27061 nuclear fusion pro
10	45	51.1	589	2	T50385 hypothetical colle
11	45	51.1	631	2	D82218 methyl-accepting c
12	45	51.1	668	2	B86989 hypothetical prote
13	45	51.1	819	2	T19351 dynein alpha heavy
14	45	51.1	2405	2	T08164 tropomyosin TPM1 -
15	44	50.0	199	2	A32183 M-related protein
16	44	50.0	304	2	F72346 phospho-beta-gluc
17	44	50.0	410	2	D64158 hypothetical prote
18	44	50.0	451	2	AE0708 hypothetical prote
19	44	50.0	461	2	AE2251 hypothetical prote
20	44	50.0	926	2	T24923 phase-related prote
21	44	50.0	1585	2	B69948 M protein precursor
22	43	48.9	74	2	SA1179 hypothetical prote
23	43	48.9	156	2	CA4824 hypothetical prote
24	43	48.9	228	2	D69770 ABC transporter (b
25	43	48.9	272	2	G69946 Ig V-region-Like B
26	43	48.9	322	2	G69946 two-component hyb
27	43	48.9	372	2	CA2415 lysine-tRNA ligase
28	43	48.9	443	2	SA2415
29	43	48.9	488	2	S69892

30	43	48.9	560	2	F71452 hypothetical prote
31	43	48.9	626	2	AG1256 DNA primase [impor
32	43	48.9	879	2	C71083 conserved hypochet
33	43	48.9	1013	2	G82450 probable exonuclea
34	43	48.9	1112	2	T40382 dna repair protein
35	43	48.9	1133	2	S30301 excision repair pr
36	42	47.7	125	2	JC4799 basic leucine zipp
37	42	47.7	150	2	A05173 thyroid hormone-in
38	42	47.7	181	2	D97763 hypothetical prote
39	42	47.7	222	2	C84461 En/Spw-like transp
40	42	47.7	314	2	T07889 CCAAT-binding fact
41	42	47.7	320	2	T41335 hypothetical prote
42	42	47.7	321	2	S49369 mobilization prote
43	42	47.7	551	2	H83458 conserved hypochet
44	42	47.7	651	2	AG0448 probable membrane
45	42	47.7	677	2	S73798 MG260 homolog H91
46	42	47.7	682	2	T40270 hypothetical prote
47	42	47.7	690	2	S41009 protein T0565.9 [i
48	42	47.7	705	2	B88564 spectrin alpha 11
49	42	47.7	819	2	S61217 protein K02A.3 [i
50	42	47.7	1020	2	G88208 spectrin alpha cha
51	42	47.7	1030	2	A32612 calcium-channel ho
52	42	47.7	1053	2	S58883 transcription fact
53	42	47.7	1093	2	A47212 NF-180 - sea lamp
54	42	47.7	1110	2	I51116 Smc protein essent
55	42	47.7	1186	2	AD1300 hypothetical prote
56	42	47.7	1186	2	AD1672 hypothetical prote
57	42	47.7	1325	2	T14790 hypothetical prote
58	42	47.7	1430	2	T34516 hypothetical prote
59	42	47.7	1529	2	T16779 fodrin alpha chain
60	42	47.7	2472	2	A35715 sensory transducti
61	41.5	47.2	1671	2	S71628 protein gp31 - pha
62	41	46.6	76	2	T13117 hypothetical prote
63	41	46.6	124	2	B97153 transcription regu
64	41	46.6	128	2	A12766 probable transcript
65	41	46.6	158	2	D97547 uncharacterized co
66	41	46.6	175	2	B97292 conserved hypochet
67	41	46.6	219	2	A64552 hypothetical prote
68	41	46.6	226	2	C82209 repetitive protein
69	41	46.6	259	2	D60110 hypothetical prote
70	41	46.6	286	2	A70168 gliding motility p
71	41	46.6	298	2	T44443 hypothetical prote
72	41	46.6	299	2	T05866 probable periplasm
73	41	46.6	300	2	B81312 D-lactatase-diophos
74	41	46.6	422	2	AB1154 probable reverse t
75	41	46.6	505	2	H95380 antigen 332 - mala
76	41	46.6	727	2	B84847 conserved hypochet
77	41	46.6	837	2	JN0292 B-G antigen - chic
78	40	45.5	83	2	H82356 B-G antigen - chic
79	40	45.5	152	2	S17935 hypothetical prote
80	40	45.5	169	2	I50166 B-G antigen - chic
81	40	45.5	200	2	H86812 B-G antigen - chic
82	40	45.5	204	2	I50163 scaffold protein -
83	40	45.5	214	2	S58140 oleosin (clone P24
84	40	45.5	223	1	S17936 oleosin P24, isofo
85	40	45.5	226	2	T06378 hypothetical prote
86	40	45.5	290	2	T22161 RNA polymerase ECF
87	40	45.5	320	2	H84103 secretory carrier
88	40	45.5	347	2	T08826 Ig V-region-Like B
89	40	45.5	432	2	A39371 transcrption init
90	40	45.5	466	1	S61292 serine/threonine k
91	40	45.5	480	2	T47255 protein kinase pkn
92	40	45.5	505	1	S77034 serine/threonine k
93	40	45.5	598	2	T47254 hypothetical prote
94	40	45.5	610	2	T16194 DNA primase [impor
95	40	45.5	626	2	AC1619 hypothetical prote
96	40	45.5	652	2	B59102 ABC transporter, p
97	40	45.5	693	2	D90441 ABC transporter, p
98	40	45.5	717	2	T49238 hypothetical prote
99	40	45.5	767	2	S63220 probable membrane
100	40	45.5	780	2	T31548 hypothetical prote

ALIGNMENTS

RESULT 1
B81192
hemagglutinin/hemolysin-related protein NMB0497 [imported] - Neisseria meningitidis (str
C/Species: Neisseria meningitidis
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C/Accession: B81192
R/Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
rt, H.; Qin, H.; Vamathavan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: B81192
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1975 <TET>
A:Cross-references: UNIPROT:O9K0S7; GB:AE002406; GB:AE002098; NID:g7225720; PIDN:AAF4092
C:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0497

Query Match 100.0%; Score 88; DB 2; Length 1975;
Best Local Similarity 100.0%; Pred. No. 7,9e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQQ 18
Db 959 PTOKAAELNOKSKELEQQ 973

RESULT 2
G81044
hemagglutinin/hemolysin-related protein NMB1779 [imported] - Neisseria meningitidis (str
C/Species: Neisseria meningitidis
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C/Accession: G81044
R/Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
rt, H.; Qin, H.; Vamathavan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: G81044
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1995 <TET>
A:Cross-references: UNIPROT:O9JY23; GB:AE002527; GB:AE002098; NID:g7227203; PIDN:AAF4211
C:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1779

Query Match 100.0%; Score 88; DB 2; Length 1995;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQQ 18
Db 959 PTOKAAELNOKSKELEQQ 976

RESULT 3
B81989
hypothetical protein NMA0688 [imported] - Neisseria meningitidis (strain Z2491 serogroup
C/Species: Neisseria meningitidis
C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C/Accession: B81989
R/Packhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagers, K.; Leather, S.; Mout, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:2022556; PMID:10761919
A:Accession: B81989
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-2015 <PAR>
A:Cross-references: UNIPROT:O9GRD2; GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83974
C:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA0688

Query Match 100.0%; Score 88; DB 2; Length 2015;
Best Local Similarity 100.0%; Pred. No. 8,1e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQQ 18
Db 959 PTOKAAELNOKSKELEQQ 976

RESULT 4
H36812
hypothetical protein ORF64 - saimiriine herpesvirus 1 (strain 11)
C/Species: saimiriine herpesvirus 1
A:Note: host Saimiri sciureus (common squirrel monkey)
C/Date: 16-Oct-1992 #sequence_revision 16-Oct-1992 #text_change 08-Oct-1999
C/Accession: H36812
R/Albrecht, J.
submitted to the EMBL Data Library, January 1992
A:Description: Primary structure of the herpesvirus saimiri genome.
A:Reference number: A36806
A:Accession: H36812
A:Molecule type: DNA
A:Residues: 1-2469 <ALB>
A:Cross-references: GB:X64346; NID:g60320; PIDN:CAA45687.1; PID:g60385
R/Albrecht, J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Blesinger, B.; Newman, C.; Wit
J. Virol. 66, 5047-5058, 1992
A:Title: Primary structure of the herpesvirus saimiri genome.
A:Reference number: A37309; MUID:92333688; PMID:1331267
A:Contents: annotation; protein-coding frames
A:Note: neither protein nor nucleotide sequence is given
C:Genetics:
A:Gene: 64

Query Match 55.7%; Score 49; DB 2; Length 2469;
Best Local Similarity 50.0%; Pred. No. 58;
Matches 9; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PTOKAAELNOKSKELEQQ 18
Db 670 PSKKAQQLHESKLVLEQK 687

RESULT 5
S19063
hypothetical protein YNR023w precursor, mitochondrial - yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein N3224
C/Species: Saccharomyces cerevisiae
C/Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 09-Jul-2004
C/Accession: S19063; S63354
R/Donald, K.A.G.; Hill, J.; Griffiths, D.E.
submitted to the EMBL Data Library, October 1991
A:Reference number: S19063
A:Accession: S19063
A:Molecule type: DNA
A:Residues: 1-566 <DON>
A:Cross-references: UNIPROT:P53628; EMBL:X62430; NID:g4106; PID:g4107
R/Pohl, T.M.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63346
A:Accession: S63354
A:Molecule type: DNA

A:Residues: 1-566 <POH>
A:Cross-references: EMBL:Z71638; NID:q1302511; PID:e239821; PID:q1302512; MIPS:YNR023W
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:SNF12
A:Cross-references: SGD:S0005306; MIPS:YNR023W
A:Map position: 14R
A:Genome: nuclear
C:Keywords: mitochondrion

Query Match 53.4%; Score 47; DB 2; Length 566;
Best Local Similarity 69.2%; Pred. No. 28;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 5 AAEINOKSKLEIQ 17
D 463 AAEINARELEIQ 475

RESULT 6

ABC transporter, phosphate-binding protein phnd [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
A:Accession: AF2084
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; Iriyuchi, Nakazaki, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF2084
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-333 <KUR>
A:Cross-references: UNIPROT:Q8YUV3; GB:BA000019; PIDN:BA073927.1; PID:q17131319; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: phnd

Query Match 52.3%; Score 46; DB 2; Length 333;
Best Local Similarity 64.3%; Pred. No. 23;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 3 OKAEINOKSKLEIQ 16
D 320 OKAEINOKSKLEIQ 333

RESULT 7

probable membrane protein YDL148c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D1566
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
A:Accession: S67696
R:Petra, J.; Blugeon, C.; Delavau, T.; Jacq, C.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67693
A:Accession: S67696
A:Molecule type: DNA
A:Residues: 1-810 <PER>
A:Cross-references: UNIPROT:Q99Y07; EMBL:Z74196; NID:q1431232; PID:e253061; PID:q1431233
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:NOP14; MIPS:YDL148C
A:Cross-references: SGD:S0002307
A:Map position: 4L
C:Keywords: transmembrane protein
F:492-508/Domain: transmembrane #status predicted <TM1>
F:561-577/Domain: transmembrane #status predicted <TM2>

Query Match 52.3%; Score 46; DB 2; Length 810;

Best Local Similarity 56.2%; Pred. No. 55;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
OY 3 OKAEINOKSKLEIQ 18
D 294 EKNAEAEKRELEIQ 309

RESULT 8

hypothetical protein TM0850 - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
A:Accession: C72327
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: C72327
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-172 <AEH>
A:Cross-references: UNIPROT:Q9WZV4; GB:AE001751; GB:AE000512; NID:q4981371; PIDN:AA03593;
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0850
C:Superfamily: heat shock protein gpH

Query Match 51.1%; Score 45; DB 2; Length 172;
Best Local Similarity 56.2%; Pred. No. 17;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
OY 2 TOKAEINOKSKLEIQ 17
D 10 TOKAEINOKSKLEIQ 25

RESULT 9

nuclear fusion protein BIK1 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YCL029C; protein YCL182
C:Species: Saccharomyces cerevisiae
C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 09-Jul-2004
A:Accession: B27061; S17472; S19356
R:Trueheart, J.; Boeke, J.D.; Fink, G.R.
Mol. Cell. Biol. 7, 2316-2328, 1987
A:Title: Two genes required for cell fusion during yeast conjugation: evidence for a phe
A:Reference number: A93093; MUID:87286864; PMID:3302672
A:Accession: B27061
A:Molecule type: DNA
A:Residues: 1-440 <TRU>
A:Cross-references: UNIPROT:P11709; EMBL:M16717; NID:q171534; PIDN:AAA34614.1; PID:q1715
R:Ramezani Rad, M.; Lutzenkirchen, K.; Xu, G.; Kleinhaus, U.; Hollenberg, C.P.
Yeast 7, 533-538, 1991
A:Title: The complete sequence of a 11,953 bp fragment from CIG on chromosome III encomp
A:Reference number: S17471; MUID:91377317; PMID:1897318
A:Accession: S17472
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-440 <RAM>
A:Cross-references: EMBL:X59720; NID:q1907116; PIDN:CAA42356.1; PID:q5335
R:Hollenberg, C.P.; Kleinhaus, U.; Lutzenkirchen, K.; Ramezani Rad, M.; Xu, G.
submitted to the Protein Sequence Database, March 1992
A:Reference number: S19350
A:Accession: S19356
A:Molecule type: DNA
A:Residues: 1-440 <HOL>
A:Cross-references: EMBL:X59720; NID:q1907116; PID:e264429; PID:q5335; MIPS:YCL029C
C:Genetics:
A:Gene: SGD:BIK1
A:Cross-references: SGD:S0000534; MIPS:YCL029C

A;Residues: 1-2405 <MIT>
A;Cross-references: EMBL:L26049; NID:g415679; PIDN:AAA57316.1; PID:g603079
A;Experimental source: strain 21gr

C/Genetics:

A/Gene: ODA11

A/Note: Intron positions not resolved (incomplete sequence)

C/Species: dynein heavy chain, ciliary

C/Keywords: nucleotide binding; P-loop

F/575-582/Region: nucleotide-binding motif A (P-loop)

Query Match 51.1%; Score 45; DB 2; Length 2405;

Best Local Similarity 62.5%; Pred. No. 2.2e+02; Mismatches 4; Indels 0; Gaps 0;

Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 OKAAELNOKSKLEBOQ 18

Db 2060 EKVALNKKVGELEBOQ 2075

RESULT 15

A32183

A/Species: yeast (*Saccharomyces cerevisiae*)

N/Alternate names: protein N2332; protein YNL079c

C/Species: *Saccharomyces cerevisiae*

C/Date: 08-Sep-1989 #sequence_revision 08-Sep-1989 #text_change 09-Jul-2004

C/Accession: A32183; S53899; S63011; S63928

R/Liu, H.; Breitscher, A.

Cell 57, 233-242, 1989

A/Title: Disruption of the single tropomyosin gene in yeast results in the disappearance

A/Reference number: A32183; MUID:89195234; PMID:2649250

A/Accession: A32183

A/Molecule type: DNA

A/Residues: 1-199 <LIU>

A/Cross-references: UNIPROT:P17536; EMBL:M25501; NID:g173037; PIDD:AAA5174.1; PID:g1730

R/Poehlmann, R.; Philippesen, P.

submitted to the EMBL Data Library, April 1995

A/Reference number: S53896

A/Accession: S53899

A/Molecule type: DNA

A/Residues: 1-199 <POE>

A/Cross-references: EMBL:X86470; NID:g791101; PIDD:CAA60179.1; PID:g791105

R/Poehlmann, R.; Philippesen, P.

submitted to the Protein Sequence Database, April 1996

A/Reference number: S62997

A/Accession: S63011

A/Molecule type: DNA

A/Residues: 1-199 <POW>

A/Cross-references: EMBL:Z71355; NID:g1301970; PIDD:CAA95953.1; PID:g1301971; MIPS:YNL07

R/Soler-Mita, A.; Saiz, J.E.; Ballesta, J.P.G.; Remacha, M.

submitted to the Protein Sequence Database, April 1996

A/Reference number: S63018

A/Accession: S63018

A/Molecule type: DNA

A/Residues: 1-199 <SOL>

A/Cross-references: EMBL:Z71355; NID:g1301970; PIDD:CAA95953.1; PID:g1301971; MIPS:YNL07

A/Experimental source: strain S288C

R/Poehlmann, R.; Philippesen, P.

Yeast 12, 391-402, 1996

A/Title: Sequencing a cosmid clone of *Saccharomyces cerevisiae* chromosome XIV reveals 12

A/Reference number: S63925; MUID:96267764; PMID:8701611

A/Accession: S63928

A/Molecule type: DNA

A/Status: nucleic acid sequence not shown; translation not shown

A/Residues: 1-199 <POB>

A/Cross-references: EMBL:X86470; NID:g791101; PIDD:CAA60179.1; PID:g791105

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995

C/Genetics:

A/Gene: SGD:TPM1

A/Cross-references: SGD:S0005023; MIPS:YNL079c

A/Map position: 14L

C/Species: *tropomyosin TPM1*

C/Keywords: coiled coil; cytoskeleton

Query Match 50.0%; Score 44; DB 2; Length 199;

Best Local Similarity 50.0%; Pred. No. 28;

Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 3 OKAAELNOKSKLEBOQ 18

Db 20 EKVELKKNKLEBOQ 35

RESULT 16

F72346

M-related protein - *Thermotoga maritima* (strain MSB8)C/Species: *Thermotoga maritima*

C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004

C/Accession: F72346

R/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

C.M.

Nature 399, 323-329, 1999

A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A/Reference number: A72200; MUID:99287316; PMID:10360571

A/Accession: F72346

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-304 <ARN>

A/Cross-references: UNIPROT:Q9WZD9; GB:AE001740; GB:AE000512; NID:g4981194; PIDD:AAD3575

A/Experimental source: strain MSB8

C/Genetics:

A/Gene: TM0671

Query Match 50.0%; Score 44; DB 2; Length 304;

Best Local Similarity 62.5%; Pred. No. 42;

Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 OKAAELNOKSKLEBOQ 18

Db 91 QKIDELSKKELEBOQ 106

RESULT 17

D64158

hypothetical protein HI0756 - *Haemophilus influenzae* (strain Rd KW20)C/Species: *Haemophilus influenzae*

C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004

C/Accession: D64158

R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A

; Goeyne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J

; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M.

Science 269, 496-512, 1995

A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A/Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.

A/Reference number: A64000; MUID:95350630; PMID:7542800

A/Accession: D64158

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-410 <RTGR>

A/Cross-references: UNIPROT:P44864; GB:U32760; GB:L42023; NID:g1573764; PIDD:AAC22415.1;

A/Note: best homolog was a hypothetical protein from *Escherichia coli*

C/Genetics:

A/Gene: SGD:TPM1

A/Cross-references: SGD:S0005023; MIPS:YNL079c

A/Map position: 14L

C/Species: *tropomyosin TPM1*

C/Keywords: coiled coil; cytoskeleton

Query Match 50.0%; Score 44; DB 2; Length 410;

Best Local Similarity 38.9%; Pred. No. 56;

Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 1 PTOKAELNOKSKLEBOQ 18

Db 27 PVSQSSDLNIOIKQIKQ 44

RESULT 18

AE0708

phospho-beta-glucosidase B [imported] - *Salmonella enterica* subsp. *enterica* serovar TyphC/Species: *Salmonella enterica* subsp. *enterica* serovar TyphA/Note: this species has also been called *Salmonella typhi*

C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C/Accession: AE0708

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, T.; Conington, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moutle, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AE0708
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-451 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02038.1; PID:G16502875; GSPDB:GN00176
C:Genetics:
A:Gene: STY1797
C:Superfamily: melibiose-specific alpha-galactosidase
Query Match 50.0%; Score 44; DB 2; Length 451;
Best Local Similarity 66.7%; Pred. No. 62;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
OY 4 KAAELNOKSKELQ 18
DB 298 KPELNVKPELEK 312
RESULT 19
AE2251
hypothetical protein all3564 [imported] - Noctoc sp. (strain PCC 7120)
C:Species: Noctoc sp. PCC 7120
A:Note: Noctoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AE2251
A:Authors: Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Matsumbe, A.; Iriduchi, R.; Kaneko, T.; Nakamura, Y.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.; Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena*
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE2251
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-461 <KUN>
A:Cross-references: UNIPROT:Q8YR84; GB:BA000019; PIDN:BA075263.1; PID:G17132697; GSPDB:G
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all3564
Query Match 50.0%; Score 44; DB 2; Length 461;
Best Local Similarity 53.3%; Pred. No. 63;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
OY 3 OKAAELNOKSKELQ 17
DB 180 ERTAELOQKMRLEE 194
RESULT 20
T24923
hypothetical protein T14G8.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24923
A:Authors: Matthews, P.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z19955
A:Accession: T24923
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-926 <MIL>
A:Cross-references: UNIPROT:Q22515; EMBL:Z67884; PIDN:CAA91809.1; GSPDB:GN00028; CESP:T1
A:Experimental source: clone T14G8
C:Genetics:
A:Gene: CESP:T14G8.3
A:Map position: X

A:Introns: 23/1; 81/3; 132/2; 162/3; 193/3; 229/3; 279/3; 339/3; 400/3; 484/3; 657/2; 83;
C:Superfamily: heat shock protein 91
Query Match 50.0%; Score 44; DB 2; Length 926;
Best Local Similarity 38.9%; Pred. No. 1.2e+02;
Matches 7; Conservative 8; Mismatches 3; Indels 0; Gaps 0;
OY 1 PTOKAAELNOKSKELQ 18
DB 824 PVKASRIAKARDLER 841
RESULT 21
B69948
phage-related protein homolog yqpo - *Bacillus subtilis*
C:Species: *Bacillus subtilis*
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: B69948
A:Authors: Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berti, R.; Kuster, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berti, C.; Bron, S.; Brouillet, S.; Brunsch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Puma, S.; Galizzi, A.; Gallier, lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y.; M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Roche, B.; Roche, B.; Rose, M.; Sadale, J.; Sekowska, A.; Seror, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:9804403; PMID:9384377
A:Accession: B69948
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1585 <KUN>
A:Cross-references: UNIPROT:P45931; GB:Z99117; GB:AL009126; NID:G2634966; PIDN:CAB14544.
A:Experimental source: strain 168
C:Genetics:
A:Gene: yqpo
Query Match 50.0%; Score 44; DB 2; Length 1585;
Best Local Similarity 47.1%; Pred. No. 2.1e+02;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
OY 2 TOKAAELNOKSKELQ 18
DB 1267 TSKASEVNAKELAKK 1283
RESULT 22
S60849
M protein precursor - *Streptococcus pyogenes* (serotype M65) (fragment)
C:Species: *Streptococcus pyogenes*
A:Variety: serotype M65
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S60849
A:Authors: A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.
Mol. Microbiol. 14, 619-631, 1994
A:Title: Non-congruent relationships between variation in *emm* gene sequences and the pop
A:Reference number: S60784; MUID:95198537; PMID:7891551
A:Accession: S60849
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-74 <MHA>
A:Cross-references: UNIPROT:Q54581; EMBL:U11980
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Superfamily: M5 protein
Query Match 48.9%; Score 43; DB 2; Length 74;
Best Local Similarity 47.1%; Pred. No. 15;

RESULT 23

hypothetical protein lmo0840 [imported] - Listeria monocytogenes (strain EGD-e)

C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AH1179

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihl, H.; Jones, L.M.; Karet, U.
D:Science 294, 849-852, 2001

A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schlueener, T.; Smoes, N.; Tietzer, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Lillie, Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1179

A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-156 <GLA>

A:Cross-references: UNIPROT:O8YG06; GB:NC_003210; PIDN:CAC98918.1; PID:gl6410228; GSPDB: C:Genetics:
A:Gene: lmo0840
C:Superfamily: hypothetical protein yhjH

Query Match 48.9%; Score 43; DB 2; Length 156;
Best Local Similarity 52.9%; Pred. No. 31;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Db 96 TTGAELNQVHDELHQE 112

RESULT 24

hypothetical protein At2g40020 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: C84824

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M. Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, C.J.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, A.; Lillie: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487; PMID:10617197
A:Accession: C84824

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-328 <STO>

A:Cross-references: UNIPROT:O8GW87; GB:AE002093; NID:g2088655; PIDN:AAB95287.1; GSPDB:GN C:Genetics:
A:Gene: At2g40020
A:Map position: 2

Query Match 48.9%; Score 43; DB 2; Length 228;
Best Local Similarity 47.1%; Pred. No. 45;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 83 PTOKAELNQSKLEEQ 99

RESULT 25

ABC transporter (binding protein) homolog ydha - Bacillus subtilis

Db 693770

C:\Species: Bacillus subtilis
C:\Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:\Accession: D69770
A:\Authors: F.; Ogaewara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertier
C.; Bron, S.; Broillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:\Authors: Foulger, D.; Foltz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallier
A.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konlingstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:\Authors: Labder, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y., M.; Ogawa, K.; Ogawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A:\Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Seliguchi, J.; Sekowska, A.; Seror
A.; Winters, P.; Wipit, A.; Yamamoto, H.; Yamane, K.; Yasumoto, A.; Tosato, V.; Uchiyama,
A:\Authors: Yoshikawa, H.F.; Zumslein, E.; Yoshikawa, H.; Danchin, A.
A:\Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:\Reference number: A69580; MUID:98044033; PMID:9384377
A:\Accession: G69946
A:\Status: preliminary; nucleic acid sequence not shown; translation not shown
A:\Molecule type: DNA
A:\Residues: 1-322 <KUN>
A:\Cross-references: UNIPROT:P45920; GB:Z99117; GB:AL009126; NID:g2634966; PIDN:CAB14556.
A:\Experimental source: strain 168
C:\Genetics:
A:\Gene: yqbd

Query Match 48.9%; Score 43; DB 2; Length 322;
Best Local Similarity 50.0%; Pred. No. 63;
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

3 OKAAELNOKSKELEQ 18

Db 254 EKEELNKKDKEKEE 269

RESULT 27

C39371
IG V-region-like B-G antigen 11/4 precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 09-Jul-2004
C:Accession: C39371
R:Miller, M.M.; Goto, R.; Young, S.; Chirivella, J.; Hawke, D.; Miyada, C.G.
Proc. Natl. Acad. Sci. U.S.A. 88, 4377-4381, 1991
A:Title: Immunoglobulin variable-region-like domains of diverse sequence within the major
A:Reference number: A39371; MUID:91239571; PMID:1903541
A:Accession: C39371
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-372 <MIL>
A:Cross-references: UNIPROT:Q31406; GB:M61861

Query Match 48.9%; Score 43; DB 2; Length 372;
Best Local Similarity 64.3%; Pred. No. 72;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 AAELNOKSKELEEQ 18
||:|||||:|
Db 270 AAKLAQCKELEKQ 283

RESULT 28

AH2415
Two-component hybrid sensor and regulator alr4880 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. strain PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH2415
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasa moto, S.; Matanabe, A.; Iriyuchi, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2415
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <KUR>
A:Cross-references: UNIPROT:Q8VWQ3; GB:BA000019; PIDN:BA876579.1; PID:gl7134017; GSPDB:C
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr4880

Query Match 48.9%; Score 43; DB 2; Length 443;
Best Local Similarity 53.3%; Pred. No. 85;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 TOKAAELNOKSKELE 16
||:|||||:|
Db 140 TERTAEINQALKDLQ 154

RESULT 29

S69892
Lysine-tRNA ligase (EC 6.1.1.6) - Mycoplasma hominis
C:Species: Mycoplasma hominis
C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 03-Jun-2002
C:Accession: S69892
R:Oekoekmen, D.; Birkelund, S.; Christiansen, G.
FEMS Microbiol. Lett. 116, 277-282, 1994
A:Title: Characterization of a Mycoplasma hominis gene encoding lysyl-tRNA synthetase (LysRS)
A:Reference number: S69892; MUID:94237425; PMID:181699
A:Accession: S69892
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-488 <OEZ>

A:Cross-references: EMBL:X74912

C:Genetics:

A:Gene: lysS

A:Genetic code: SGC3

C:Superfamily: lysine-tRNA ligase

C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

Query Match 48.9%; Score 43; DB 2; Length 488;
Best Local Similarity 44.4%; Pred. No. 94;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 PTQAAELNOKSKELEEQ 18
||:|||||:|
Db 33 PTANSKEINKEYNLSRE 50

RESULT 30

F71452
Hypothetical protein PH0277 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
C:Accession: F71452
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi, A.
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: F71452
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-560 <KAW>
A:Cross-references: UNIPROT:O58015; GB:AF000001; MID:G3236128; PIDN:BAA29349.1; PID:G325
A:Experimental source: strain OT3
C:Genetics:
A:Gene: PH0277

Query Match 48.9%; Score 43; DB 2; Length 560;
Best Local Similarity 53.3%; Pred. No. 1,1e+02;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 4 KAAELNOKSKELEEQ 18
||:|||||:|
Db 415 KAAELNKKKEIEKE 429

Search completed: December 30, 2004, 17:00:49
Job time : 42 secs